

МУЛЬТИ-ТАЙМФРЕЙМОВЫЕ DRUMMOND-ПАТЧИ И JERA-ПРЕДОБУЧЕНИЕ ДЛЯ КРАТКОСРОЧНОГО ПРОГНОЗА РОЗНИЧНЫХ ОНЛС-РЯДОВ

А. С. Сизов¹ [0000-0001-8110-9929], Ю. А. Халин² [0000-0002-7020-8515],

А. А. Белых³ [0009-0005-7408-0052]

¹⁻³Юго-западный государственный университет, г. Курск, Россия

¹kafedra-ipm@mail.ru, ²yur-khalin@yandex.ru, ³belykhartem.a@mail.ru

Аннотация

Предложен метод построения инвариантных к масштабу представлений временных рядов розничной выручки на базе трехбарной (по трем соседним периодам) геометрии Драммонда (DG), расширенной мульти-таймфреймовым контекстом (день, частичная календарная неделя и скользящая 7-дневка). На этих «патчах» выполнено self-supervised предобучение по схеме Joint-Embedding Predictive Architecture (JERA) со спATIO-темпоральным маскированием, после чего модель дообучена с выходными слоями, оценивающими неопределенность, для прогноза на следующий день и следующую неделю. Проанализированы свойства аффинной инвариантности признаков и идентифицируемости недельной фазы; эмпирически продемонстрировано улучшение по сравнению с сильными базовыми моделями на реальных данных.

Ключевые слова: геометрия Драммонда, Joint-Embedding Predictive Architecture (JERA), временные ряды, Open-High-Low-Close (ОНЛС), розничная торговля, краткосрочный прогноз, самообучение.

ВВЕДЕНИЕ

Современные розничные временные ряды характеризуются высокой вариативностью, выраженной недельной сезонностью и наличием шума, что создает существенные сложности для краткосрочного прогнозирования ключевых метрик: выручки, количества кассовых чеков и среднего чека [1, 2]. Традиционные методы, такие как ARIMA и экспоненциальное сглаживание, могут демонстрировать недостаточную точность на нелинейных зависимостях и при резких изменениях режимов [3]. Нейросетевые подходы, включая N-BEATS и Temporal

Fusion Transformer (TFT), показывают лучшие результаты, но требуют больших объемов данных и могут быть неустойчивы к изменениям масштаба и сдвигам уровня ряда [4, 5].

В настоящей работе рассмотрена комбинация классических технических индикаторов и современных методов самообучающихся представлений. В качестве основы для признаков использована геометрия Драммонда (DG) [6] – набор интерпретируемых уровней, отражающих локальную геометрию ценового движения, обобщенную на случай произвольных временных рядов. Для обучения представлений применена архитектура JERA (Joint-Embedding Predictive Architecture), показавшая свою эффективность в задачах компьютерного зрения и обработки сигналов [7, 8]. Ключевая идея JERA – предсказание представлений одних частей данных по контексту других, что позволяет модели извлекать устойчивые скрытые (латентные) зависимости без реконструкции входных данных.

Перечислим основные полученные результаты.

1. Предложена единая многошкальная по времени (мульти-таймфреймовая) постановка задачи для трех каналов розничных данных (выручка, чеки, средний чек) с использованием недельно-базисных приращений и OHLC-агрегирования, где Open – цена открытия, High – максимальная цена, Low – минимальная цена и Close – цена закрытия.

2. Разработан метод построения аффинно-инвариантных фрагментов данных (Drummond-патчей), объединяющих информацию с дневного и двухнедельных горизонтов.

3. Адаптирована и доработана схема JERA-предобучения для временных рядов с пространственно-временным маскированием и позиционным кодированием, учитывающим недельную фазу.

4. На реальных данных выполнено экспериментальное сравнение предложенного подхода с рядом сильных базовых моделей, показавшее статистически значимое улучшение качества на краткосрочных горизонтах: следующий день (D+1) и следующая неделя (W+1). Эффективность самообучения для розничных продаж, таким образом, получила дополнительное подтверждение [9].

ДАННЫЕ И ОБОЗНАЧЕНИЯ

Использованы три канала, отражающие изменение динамики базовых розничных метрик:

- GainVal – изменение выручки;
- CheckCount – изменение количества покупок (кассовых чеков);
- ARVal – изменение среднего чека (Average Receipt Value).

Пусть V_t – выручка в день t , N_t – число кассовых чеков и $A_t = V_t/N_t$ – средний чек. Для подавления сезонности по дням недели используем недельно-базисные приращения (WeekBasis) в виде лог-отношений:

$$\Delta^W X_t = \log(X_t + \varepsilon) - \log(X_{(t-7)} + \varepsilon), \text{ где } X \in \{V, N, A\}, \varepsilon > 0. \quad (1)$$

По умолчанию в тексте под именами каналов понимаем именно эти WeekBasis-приросты:

$$\text{GainVal}_t \equiv \Delta^W V_t, \quad \text{CheckCount}_t \equiv \Delta^W N_t, \quad \text{ARVal}_t \equiv \Delta^W A_t.$$

Для задач Day(t) может использоваться дневной аналог недельного приращения (1):

$$\Delta^D X_t = \log(X_t + \varepsilon) - \log(X_{t-1} + \varepsilon).$$

Дополнительно используем OHLC-агрегирование в заданном окне $[a, b]$:

$$\text{OHLC}_{[a,b]}(y) = (O, H, L, C) = \left(y_a, \max_{t \in [a,b]} y_t, \min_{t \in [a,b]} y_t, y_b \right).$$

Рассмотрим два недельных варианта окон для OHLC над приращениями $\Delta^W X$:

- Week_cal: календарная неделя (понедельник – воскресенье), содержащая t ;
- Week_roll: скользящее окно $[t - 6, t]$ из 7 дней.

На рис. 1 проиллюстрировано построение OHLC для канала GainVal; аналогично могут быть построены OHLC для каналов CheckCount и ARVal.

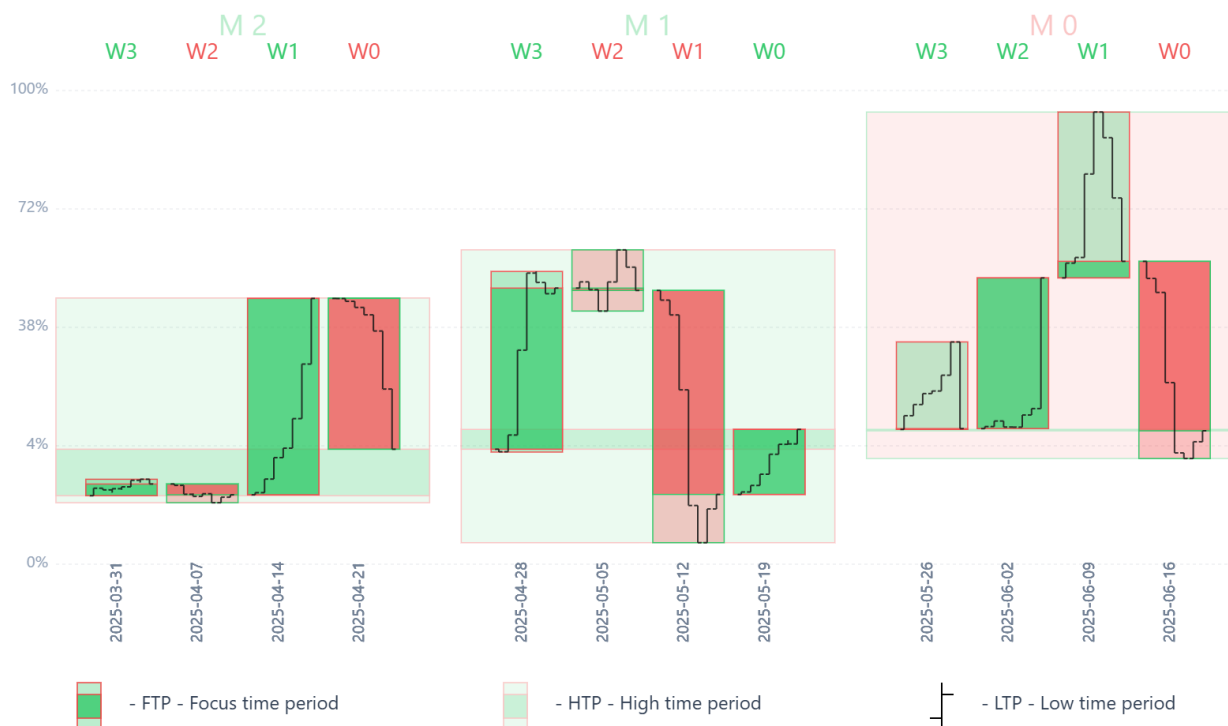


Рис. 1. Схема построения OHLC по WeekBasis-приростам выручки (GainVal).

1. МЕТОД: DRUMMOND-ПАТЧ И JERA-ПРЕДОБУЧЕНИЕ

1.1. Построение мульти-таймфреймового Drummond-патча

Геометрия Драммонда (DG) определяет набор уровней на основе цен OHLC трех последовательных временных интервалов (баров) [6]. Для окна из трех баров (H_i, L_i, C_i) , $i \in \{0,1,2\}$ (0 – текущий) определим базовые уровни:

$$\text{pivot}_i = (H_i + L_i + C_i)/3,$$

$$h_3 = (H_0 + H_1 + H_2)/3,$$

$$l_3 = (L_0 + L_1 + L_2)/3,$$

$$\text{pld} = (\text{pivot}_0 + \text{pivot}_1 + \text{pivot}_2)/3,$$

$$\text{rbird} = (\text{pivot}_0 + \text{pivot}_1 + C_0)/3.$$

На их основе вычислим производные уровни, такие как $et_1 = 2 \text{pld} - l_3$, $eb_1 = 2 \text{pld} - h_3$ и др. [6].

Патч P_t на конец дня t включает DG-уровни и z-координаты (нормализованные значения) для трех временных масштабов (тайм фреймов) τ :

- Day: последние три дня $[t - 2, t]$;
- Week_cal: две прошлые полные календарные недели + текущая частичная (с понедельника по день t) (Mon... t);
- Week_roll: три скользящих окна $[t - 20 \dots t - 14]$, $[t - 13 \dots t - 7]$, $[t - 6 \dots t]$.

Для обеспечения аффинной инвариантности используем нормализацию внутри каждого окна $[a, b]$ по High/Low для каждой компоненты j :

$$z_{[a,b]}(x_t)^{(j)} = \frac{x_t^{(j)} - L_j}{\max\{H_j - L_j, \varepsilon\}} \quad (2)$$

где $H_j = \max_{t \in [a,b]} x_t^{(j)}$, $L_j = \min_{t \in [a,b]} x_t^{(j)}$.

Лемма 1 (Аффинная инвариантность нормализации). Для любого аффинного преобразования $x \mapsto ax + b$ с $a > 0$ и любого окна $[a, b]$ выполняется $z_{[a,b]}(ax + b) = z_{[a,b]}(x)$.

Доказательство следует непосредственно из определения (2).

Патч также обогащается межмасштабными признаками (позиция дневного бара относительно недельных уровней) и календарными мета-признаками (день недели, признак незавершенной (частичной) недели). Все расчеты произведены строго на исторических данных без утечки из будущего.

1.2. JEPA-обучение для розничных OHLC-представлений

Архитектура обучения, представленная на рис. 2, следует принципам JEPA [7, 8] в контексте задач временных рядов [10].

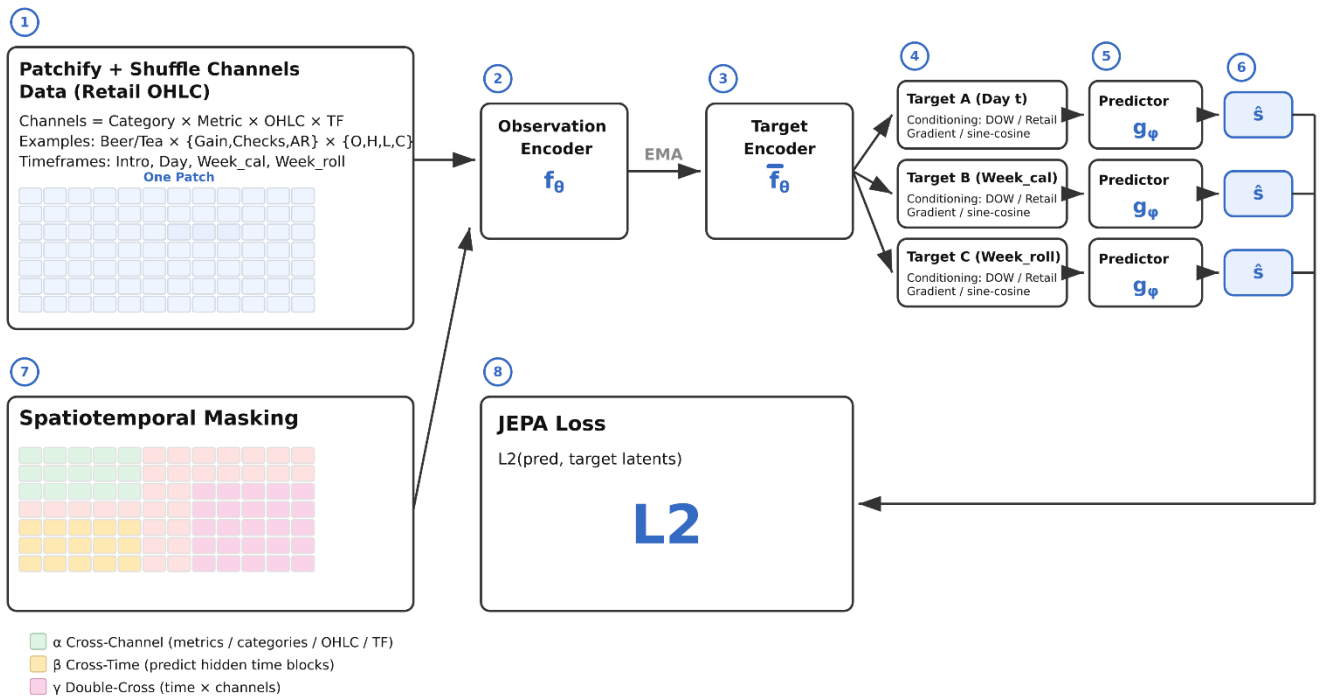


Рис. 2. Retail-JEPA: схема обучения и предсказания для мультишкальных OHLC-данных.

Пояснения к рис. 2.

Разбиение на фрагменты и перемешивание каналов (Patchify+Shuffle Channels). Вход – многоканальные розничные временные ряды формата Категория × Метрика × OHLC × Тайм фрейм.

Метрики: {GainVal, CheckCount, ARVal}; OHLC: {O, H, L, C}; тайм фреймы: {Intro, Day, Week_cal, Week_roll}. Поток разрезан на компактные DG-патчи (трехбарные блоки); каналы могут перемешиваться для регуляризации.

Кодировщик наблюдаемой части (Observation Encoder) f_θ . Незамаскированная (наблюдаемая) часть патча кодируется кодировщиком (энкодером) наблюдения f_θ в скрытое представление (латент) s_x . Во вход уже заложена DG-нормализация $(X - PLdot)/\Delta$, что делает признаки аффинно-инвариантными и снижает влияние локальных амплитуд.

EMA (экспоненциальное скользящее среднее) → Target Encoder \bar{f}_θ . Целевой кодировщик (таргет-энкодер) \bar{f}_θ есть экспоненциально сглаженная (EMA) копия параметров f_θ . Он кодирует скрытые блоки (замаскированные регионы пат-

ча) в целевые скрытые представления (таргет-латенты) s_y^τ с остановкой градиента. Это стабилизирует целевые представления и предотвращает вырождение представлений (коллапс).

Позиционное обусловливание (Positional Conditioning). К латенту s_x добавляются позиционные признаки: (а) Temporal – день недели (DOW), синус/косинус-время, флаг partial-week (неполная неделя обрезана на текущем дне t); (б) Retail Gradient Positioning – векторные представления (эмбединги) структурных осей (категория товара, тип метрики, OHLC-канал, тип таймфрейма, Week_cal vs Week_roll). Эти признаки сообщают модели фазу недели (например, пятница/суббота пик спроса) и контекст тайм фрейма.

Предсказывающие головы/модули (Predictors) g_ϕ . Небольшие прогнозирующие (предикторные) головы g_ϕ (ViT/MLP-блоки) по объединенному представлению строят оценки латентов скрытых целей: Target A: дневной блок Day(t); Target B: Week_cal (календарная неделя Mon...t, обрезанная на текущем дне); Target C: Week_roll (скользящее окно $t-6...t$). Выходы обозначены как \hat{s}_y^τ .

Предсказанные скрытые представления целей (Predicted Target Latents). \hat{s}_y^τ – предсказанные латентные представления целевых (скрытых) блоков в пространстве \bar{f}_θ . Мы восстанавливаем напрямую не сами ряды, а их латенты, что подчеркивает структуру зависимостей «день \leftrightarrow неделя» и улучшает переносимость признаков к головам прогноза $D+1/W+1$.

Пространственно-временное скрывание (Spatiotemporal Masking). Применяем комбинированную маску без доступа к будущему:

α Cross-Channel – скрываем набор каналов (например, все каналы Week_cal или часть OHLC/метрик);

β Cross-Time – скрываем целиком временной блок (например, весь день t);

γ Double-Cross – одновременное скрывание по времени и каналам (например, Day(t)+Week_cal).

Для Week_cal берем только данные Mon...t; для Week_roll – строго окно $t-6...t$, т. е. информация из будущих дней не используется.

Функция потерь JEPA (JEPA Loss). Обучение идет по L2-расхождению между предсказанными и таргет-латентами: $L_{JEPA} = \sum_\tau \| \hat{s}_y^\tau - s_y^\tau \|_2^2$. Параметры \bar{f}_θ

обновляются только через EMA, а не напрямую градиентом, что стабилизирует целевое пространство.

Пусть $P_t = \{p_t^{(k)}\}_{k=1}^K$ – набор патчей вокруг момента t . Применим двоичную маску $M \subseteq \{1, \dots, K\}$, разделяющую P_t на контекст (наблюдаемую часть) $P_t^{\setminus M}$ и таргет (замаскированную часть) P_t^M .

Контекстный энкодер f_θ преобразует наблюдаемую часть в латентное представление $s_x = f_\theta(P_t^{\setminus M})$. Целевой энкодер \bar{f}_θ , являющийся экспоненциально сглаженной (EMA) копией f_θ , обрабатывает исходный, незамаскированный патч P_t и извлекает эталонные представления $s_y^\tau = \bar{f}_\theta(P_t^M)$ для тех его частей, которые в данном примере обучения соответствуют маске M и тайм фрейму τ . Градиент через \bar{f}_θ не пропускается (stop-gradient).

Предиктор g_ϕ , получая на вход s_x и позиционные признаки (день недели, тип тайм фрейма), предсказывает латенты целевых патчей: $\hat{s}_y^\tau = g_\phi(s_x, \tau)$.

Целевая функция обучения – это минимизация L2-расстояния между предсказанными и эталонными скрытыми представлениями (латентными векторами):

$$L_{\text{JERA}}(\theta, \phi) = \mathbb{E} [\| g_\phi(f_\theta(P_t^{\setminus M})) - \text{sg}(\bar{f}_\theta(P_t^M)) \|_2^2],$$

где P_t^M обозначает часть исходного фрагмента данных (патча), соответствующую схеме маскирования M (какие части входа скрываются), математическое ожидание берем по t , маскам M и весам w_τ .

В соответствии с теорией оптимального прогнозирования, это оптимальный предиктор в таком сценарии стремится к условному математическому ожиданию целевых представлений при данном контексте.

2. ЭКСПЕРИМЕНТЫ

2.1. Данные и настройка эксперимента

Эксперименты проводились на реальных данных розничной сети за период с 2020 по 2025 г. Были использованы данные по двум товарным категориям: пиво (Beer) и чай/кофе/какао (TeaCoffeeCocoa). Прогноз строился для трех каналов: GainVal, CheckCount, ARVal. Выборка была разделена на обучающую (2020–

2024), проверочную (I квартал 2025 г.) и тестовую (II квартал 2025 г.) части с соблюдением временного порядка.

2.2. Детали реализации

Энкодер f_θ и предиктор g_ϕ были реализованы на основе трансформерной архитектуры с 4 слоями, 8 «головами» механизма внимания (параллельными каналами attention) и размерностью скрытого состояния 256. Размерность латентного представления $s = 128$. Вероятность маскирования патча – 30%. Коэффициент EMA ($\bar{\theta}$) для целевого энкодера – 0.99. Обучение проводилось оптимизатором AdamW со скоростью обучения (learning rate) 10^{-4} и размером мини-пакета данных (батча) 128 в течение 100 эпох. Предобучение JERA заняло приблизительно два дня на графическом ускорителе GPU NVIDIA V100. После предобучения к латентным представлениям добавлялись простые выходные слои, оценивающие распределение будущих значений (два полносвязных слоя), и модель дообучалась на задаче прогнозирования.

2.3. Модели и метрики

Предложенный метод (JERA+Heads) сравнивался со следующими бэйзлайнами:

- SeasonalNaive: наивный прогноз с недельной сезонностью;
- Ridge/LightGBM: линейная модель и градиентное усиление (бустинг) на табличных признаках, лежащие в основе современных ансамблевых решений [11];
- N-BEATS/N-HiTS: современные нейросетевые модели для прогнозирования временных рядов [4];
- TFT: Temporal Fusion Transformer [5].

С целью оценки вкладов компонентов (ablation study) были также протестированы:

- LightGBM (DG-признаки): LightGBM, обученный на сконструированных Drummond-патчах;
- TFT (DG-признаки): модель TFT, обученная на тех же патчах с DG-признаками сквозным образом (end-to-end);

- JEPA+Heads (Raw Features): наш метод, но на сырых недельных приращениях без DG-обработки.

Целями прогнозирования были:

- D+1: прогноз дневного приращения $\Delta^D X_{(t+1)}$;
- W+1: прогноз OHLC по $\Delta^W X$ на следующей неделе.

Использовались следующие метрики: симметричная средняя абсолютная процентная ошибка (sMAPE), средняя абсолютная масштабированная ошибка (MASE), для оценки статистической значимости различий применялся тест Диболда – Мариано [12].

2.4. Результаты и их обсуждение

Основные результаты, усредненные по всем каналам и категориям, представлены в табл. 1.

Табл. 1. Основные результаты на GainVal/CheckCount/ARVal: усредненные ошибки (меньше – лучше).

Модель	sMAPE (D+1)	sMAPE (W+1)	MASE (D+1)	MASE (W+1)
SeasonalNaive	21.8	24.9	1.00	1.00
Ridge / LightGBM	19.6	22.3	0.92	0.95
N-BEATS / N-HiTS	18.0	20.5	0.86	0.90
TFT	17.2	19.9	0.84	0.88
JEPA+Heads (предл.)	14.9	17.6	0.76	0.82

Предложенный метод JEPA+Heads показал наилучшие результаты по всем метрикам на обоих горизонтах прогноза. Улучшение по сравнению с лучшим из бэйзлайнов (TFT) составило около 13% по sMAPE на горизонте D+1 и 11%

на горизонте W+1. Результаты теста Диболда – Мариано подтвердили статистическую значимость улучшений (значение p-value против модели TFT составило 0.003 для D+1 и 0.007 для W+1). Анализ разбивки результатов по отдельным каналам и категориям показал согласованные улучшения, с максимальным выигрышем на канале GainVal (15.2% на D+1).

Результаты ablation study (табл. 2) показывают вклад каждого компонента метода. LightGBM на DG-признаках уже демонстрирует улучшение над LightGBM на сырых данных, что подтверждает полезность самих Drummond-патчей. TFT, обученная на DG-признаках, показывает результат, близкий к оригинальной TFT, что говорит о сложности прямого использования этих признаков без специального предобучения. Наш метод без DG-признаков (JEPA+Heads (Raw Features)) уступает полной модели, но все же превосходит TFT, что доказывает эффективность JEPA-предобучения. Наилучший результат достигается только при совместном использовании DG-признаков и JEPA-предобучения.

Табл. 2. Исследование методом абляции: sMAPE на горизонте D+1 (усреднено).

Модель	sMAPE (D+1)
LightGBM (Raw Features)	19.6
LightGBM (DG-признаки)	18.1
TFT (Raw Features)	17.2
TFT (DG-признаки)	17.4
JEPA+Heads (Raw Features)	16.0
JEPA+Heads (DG-признаки, полный метод)	14.9

Эффективность предложенного метода объясняется сочетанием аффинно-инвариантных DG-признаков, которые устойчивы к изменениям масштаба ряда, и JEPA-предобучения, которое позволяет извлекать информативные представления, согласованные между различными временными горизонтами.

ЗАКЛЮЧЕНИЕ

Представлен метод прогнозирования розничных временных рядов, сочетающий построение мульти-таймфреймовых Drummond-патчей и self-supervised предобучение по схеме JERA, когда целевые сигналы формируются автоматически из исходных данных. Ключевыми особенностями метода являются аффинно-инвариантная нормализация признаков, пространственно-временное маскирование патчей и использование EMA-таргет энкодера для стабилизации обучения.

Эксперименты на реальных данных показали, что предложенный подход статистически значительно превосходит сильные бэйзлайны на горизонтах прогноза D+1 и W+1. Исследование методом абляции подтвердило важность каждого компонента метода. Полученные результаты свидетельствуют о перспективности метода для практического применения в задачах операционного планирования в розничной торговле.

Основными ограничениями работы являются локальность трехбарного анализа DG, зависимость от схемы маскирования и отсутствие учета внешних факторов (праздники, акции). Перспективными направлениями дальнейших исследований являются интеграция внешних признаков, разработка более сложных стратегий маскирования и масштабирование метода на большее число товарных категорий.

СПИСОК ЛИТЕРАТУРЫ

1. *Fildes R., Ma S., Kolassa S.* Retail forecasting: Research and practice // International Journal of Forecasting. 2022. Vol. 38, No. 4. P. 1283–1318. <https://doi.org/10.1016/j.ijforecast.2019.06.004>
2. *Lim B., Arik S. O., Loeff N., Pfister T.* Temporal Fusion Transformers for interpretable multi-horizon time series forecasting // International Journal of Forecasting. 2021. Vol. 37, No. 4. P. 1748–1764. <https://doi.org/10.1016/j.ijforecast.2021.03.012>
3. *Hyndman R. J., Athanasopoulos G.* Forecasting: Principles and Practice. 2nd ed. Melbourne: OTexts, 2018. 380 p. Цит. по с.: 183–220, 221–274, 347–368.
4. *Oreshkin B.N., Carпов D., Chapados N., Bengio Y.* N-BEATS: Neural basis expansion analysis for interpretable time series forecasting // arXiv preprint

arXiv:1905.10437. 2019. <https://doi.org/10.48550/arXiv.1905.10437>; Challu C. et al. NHITS: Neural hierarchical interpolation for time series forecasting // Proceedings of the AAAI Conference on Artificial Intelligence. 2023. Vol. 37, No. 6. P. 6989–6997. <https://doi.org/10.1609/aaai.v37i6.25854>

5. Yue Zh. et al. TS2Vec: Towards Universal Representation of Time Series // Proceedings of the AAAI Conference on Artificial Intelligence. 2022. Vol. 36, No. 8. P. 8980–8987. <https://doi.org/10.1609/aaai.v36i8.20881>

6. Hearne T. Drummond Geometry: Picking Yearly Highs and Lows in Inter-bank Forex Trading // Breakthroughs in Technical Analysis: New Thinking from the World's Top Minds / ed. by D. Keller. Princeton: Bloomberg Press, 2007. P. 1–19. <https://doi.org/10.1002/9781119204749.ch1>.

7. Dawid A., LeCun Y. Introduction to Latent Variable Energy-Based Models: A Path Towards Autonomous Machine Intelligence // Journal of Statistical Mechanics: Theory and Experiment. 2024. No. 10. Art. 104011. <https://doi.org/10.1088/1742-5468/ad292b>. (arXiv:2306.02572)

8. Assran M. et al. Self-supervised learning from images with a joint-embedding predictive architecture // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023. P. 15619–15629. <https://doi.org/10.1109/CVPR52729.2023.01499>

9. Park Y.J. et al. A scalable and transferable time series prediction framework for demand forecasting. 2024. arXiv preprint arXiv:2402.19402. <https://doi.org/10.48550/arXiv.2402.19402>

10. Ragab M., Liu Q., Jia W., Chen M., Yun U. Self-Supervised Learning for Time Series Analysis: Taxonomy, Progress, and Prospects // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2024. Vol. 46, No. 10. P. 6775–6794. <https://doi.org/10.1109/TPAMI.2024.3387317>

11. Волошин Т.А., Зайцев К.С., Дунаев М.Е. Применение адаптивных ансамблей методов машинного обучения к задаче прогнозирования временных рядов // International Journal of Open Information Technologies. 2023. Т. 11, №. 8. С. 57–63.

12. Diebold F.X., Mariano R.S. Comparing Predictive Accuracy // Journal of Business & Economic Statistics. 1995. Vol. 13, No. 3. P. 253–263.
<https://doi.org/10.1080/07350015.1995.10524599>

MULTI-TIMEFRAME DRUMMOND PATCHES AND JEP A PRE-TRAINING FOR SHORT-TERM RETAIL OHLC SERIES FORECASTING

A. S. Sizov¹ [0000-0001-8110-9929], Y. A. Khalin² [0000-0002-7020-8515],
A. A. Belykh³ [0009-0005-7408-0052]

^{1–3}Southwest State University, Kursk, Russia

¹kafedra-ipm@mail.ru, ²yur-khalin@yandex.ru, ³belykhartem.a@mail.ru

Abstract

We propose a method for constructing scale-invariant representations of retail revenue time series based on three-bar Drummond Geometry (DG) computed over three adjacent periods, extended with a multi-timeframe context (day, partial calendar week, and a rolling 7-day window). Self-supervised pre-training on these “patches” is performed using a Joint-Embedding Predictive Architecture (JEP A) with spatio-temporal masking, followed by fine-tuning with output heads that quantify predictive uncertainty for next-day and next-week forecasts. The work analyzes the properties of affine invariance of the features and the identifiability of the weekly phase; empirical improvement over strong baseline models on real-world data is demonstrated.

Keywords: Drummond Geometry, Joint-Embedding Predictive Architecture (JEP A), time series, Open-High-Low-Close (OHLC), retail, short-term forecasting, self-supervised learning.

REFERENCES

1. Fildes R., Ma S., Kolassa S. Retail forecasting: Research and practice // International Journal of Forecasting. 2022. Vol. 38, No. 4. P. 1283–1318.
<https://doi.org/10.1016/j.ijforecast.2019.06.004>

2. *Lim B., Arik S.O., Loeff N., Pfister T.* Temporal Fusion Transformers for interpretable multi-horizon time series forecasting // *International Journal of Forecasting*. 2021. Vol. 37, No. 4. P. 1748–1764.
<https://doi.org/10.1016/j.ijforecast.2021.03.012>
3. *Hyndman R.J., Athanasopoulos G.* *Forecasting: Principles and Practice*. 2nd ed. Melbourne: OTexts, 2018. 380 p. Cited pp.: 183–220, 221–274, 347–368.
4. *Oreshkin B.N., Carpov D., Chapados N., Bengio Y.* N-BEATS: Neural basis expansion analysis for interpretable time series forecasting // *arXiv preprint arXiv:1905.10437*. 2019. <https://doi.org/10.48550/arXiv.1905.10437>; *Challu C. et al.* NHITS: Neural hierarchical interpolation for time series forecasting // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2023. Vol. 37, No. 6. P. 6989–6997.
<https://doi.org/10.1609/aaai.v37i6.25854>
5. *Yue Zh. et al.* TS2Vec: Towards Universal Representation of Time Series // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2022. Vol. 36, No. 8. P. 8980–8987. <https://doi.org/10.1609/aaai.v36i8.20881>
6. *Hearne T.* Drummond Geometry: Picking Yearly Highs and Lows in Interbank Forex Trading // *Breakthroughs in Technical Analysis: New Thinking from the World's Top Minds* / ed. by D. Keller. Princeton: Bloomberg Press, 2007. P. 1–19.
<https://doi.org/10.1002/9781119204749.ch1>
7. *Dawid A., LeCun Y.* Introduction to Latent Variable Energy-Based Models: A Path Towards Autonomous Machine Intelligence // *Journal of Statistical Mechanics: Theory and Experiment*. 2024. No. 10. Art. 104011.
<https://doi.org/10.1088/1742-5468/ad292b> (arXiv:2306.02572).
8. *Assran M. et al.* Self-supervised learning from images with a joint-embedding predictive architecture // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2023. P. 15619–15629.
<https://doi.org/10.1109/CVPR52729.2023.01499>
9. *Park Y.J. et al.* A scalable and transferable time series prediction framework for demand forecasting. 2024. *arXiv preprint arXiv:2402.19402*.
<https://doi.org/10.48550/arXiv.2402.19402>

10. *Ragab M., Liu Q., Jia W., Chen M., Yun U.* Self-Supervised Learning for Time Series Analysis: Taxonomy, Progress, and Prospects // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2024. Vol. 46, No. 10. P. 6775–6794.

<https://doi.org/10.1109/TPAMI.2024.3387317>

11. *Voloshin T.A., Zaitsev K.S., Dunaev M.E.* Primenenie adaptivnykh ansamblei metodov mashinnogo obucheniya k zadache prognozirovaniya vremennykh ryadov [Application of adaptive ensembles of machine learning methods to the problem of time series forecasting] // *International Journal of Open Information Technologies*. 2023. Vol. 11, No. 8. P. 57–63.

12. *Diebold F.X., Mariano R.S.* Comparing Predictive Accuracy // *Journal of Business & Economic Statistics*. 1995. Vol. 13, No. 3. P. 253–263.

<https://doi.org/10.1080/07350015.1995.10524599>

СВЕДЕНИЯ ОБ АВТОРАХ



СИЗОВ Александр Семенович – доктор технических наук, профессор, профессор кафедры «Программная инженерия», Юго-Западный государственный университет.

Alexander Semenovich SIZOV – Doctor of technical sciences. (Engineering), Professor, Professor at the Software Engineering Department, Southwest State University.

email: kafedra-ipm@mail.ru

ORCID: 0000-0001-8110-9929



ХАЛИН ЮРИЙ Алексеевич – кандидат технических наук, доцент, доцент кафедры «Программная инженерия», Юго-Западный государственный университет.

Yuri Alekseevich KHALIN – Cand. of Sci. (Engineering), Associate Professor, Associate Professor at the Software Engineering Department, Southwest State University.

email: yur-khalin@yandex.ru

ORCID: 0000-0002-7020-8515



БЕЛЫХ Артем Александрович – аспирант, Юго-Западный государственный университет.

Artem Aleksandrovich BELYKH – Postgraduate Student, Southwest State University.

email: belykhartem.a@mail.ru

ORCID: 0009-0005-7408-0052

Материал поступил в редакцию 18 декабря 2025 года