Символьные, численные и свёрточные методы анализа временных рядов: качество и время вычислений

Г. В. Журавлёв*, М. И. Кумсков‡

На восьми наборах UCR сопоставлены семь классификаторов временных рядов: ROCKET, MiniRocket и BOSS вошли в верхнюю группу по средним рангам (1,63, 1,81 и 2,56) без различий внутри неё по критерию Неменьи, а суммарное время MiniRocket и BOSS в CPU-прогонах составило 2,4 мин и 11,6 ч. Пять моделей прогнозирования цен BTC и ETH имели больший RMSE, чем Naive, во всех восьми сочетаниях актива и горизонта.

Ключевые слова: временные ряды, классификация, символьное представление, BOSS, ROCKET, MiniRocket, вычислительная сложность, прогнозирование.

1. Введение

В словарных методах окна ряда переводятся в символьные слова, по встречаемости которых строятся признаки (BOSS [1]); методы на случайных свёртках (ROCKET [2], MiniRocket [3]) отображают ряд набором свёрточных признаков с линейным классификатором. Цель работы — на восьми выбранных наборах UCR сопоставить качество семи классификаторов и измерить время полного цикла BOSS, ROCKET и MiniRocket в проведённых прогонах, а также сравнить по RMSE пять моделей многошагового прогнозирования цен BTC и ETH с прогнозом Naive.

* *Журавлёв Георгий Валерьевич* — студент кафедры вычислительной математики механико-математического факультета МГУ имени М. В. Ломоносова, e-mail: georgii.zhuravlev@math.msu.ru.

Zhuravlev Georgy Valerievich — Student, Department of Computational Mathematics, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University.

‡ *Кумсков Михаил Иванович* — д.ф.-м.н., заведующий лабораторией вычислительных методов, профессор кафедры вычислительной математики механико-математического факультета МГУ имени М. В. Ломоносова, e-mail: mikhail.kumskov@math.msu.ru.

Kumskov Mikhail Ivanovich — Doctor of Physical and Mathematical Sciences, Head of the Laboratory of Computational Methods, Professor, Department of Computational Mathematics, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University.

Опубликовано на условиях лицензии CC-BY 4.0
Published under CC-BY 4.0

2. Данные и протокол

Классификация. Восемь наборов UCR [4] (23–8926 обучающих рядов, стандартные разбиения), не случайная выборка архива. BOSS, ROCKET и MiniRocket — классы `BOSSEnsemble`, `RocketClassifier` и `MiniRocketClassifier` библиотеки aeon (версия не зафиксирована) с параметрами по умолчанию и `random_state=42`, один запуск. SAX-BoW — мешок SAX-слов (pyts; окно 4, слово 4, алфавит 8) с логистической регрессией ($C = 1$); SAX+Tr. — SAX-код ряда (16 сегментов, алфавит 8), эмбеддинг размерности 32 и Transformer-энкодер; PatchTST — собственная реализация для классификации (патчи 16, шаг 8). Оба энкодера: $d = 64$, 4 головы, 2 слоя, AdamW (шаг 10^{-3}), 30–50 эпох с ранней остановкой по 15 % обучающей выборки; для этих трёх методов accuracy усреднена по 2–3 значениям seed. Ранги сравниваются тестом Фридмана с критерием Неменьи [5]. Время цикла «обучение + предсказание» измерено на CPU сред Kaggle и Colab без фиксации среды и числа потоков каждого запуска, поэтому отношения времён ориентировочные.

Прогнозирование. Часовые цены закрытия BTCUSDT и ETHUSDT (Binance) за 730 дней до выгрузки; дата выгрузки не сохранилась, и тестовые участки разных моделей могут быть сдвинуты на время между запусками. Модели DLinear, TiDE, TimesNet, PatchTST и NHITS (neuralforecast 1.7.5, параметры по умолчанию, вход 168 ч, 600 шагов, один запуск) обучаются на первых 85 % ряда логарифмических доходностей r и без переобучения прогнозируют неперекрывающиеся окна последних 15 % с горизонтами $H \in \{5, 10, 15, 24\}$. Прогноз цены на k-м шаге окна равен $P_0 \exp(\hat{r}_1 + \cdots + \hat{r}_k)$, где P_0 — последняя известная цена; Naive повторяет P_0.

3. Результаты

Тест Фридмана выявляет различия рангов семи методов ($\chi^2 = 42{,}1$, $p = 1{,}7 \cdot 10^{-7}$). Лучшие средние ранги у ROCKET, MiniRocket и BOSS (табл. 1); их попарные разности не превышают 0,94 при критической разности Неменьи CD = 3,185 ($\alpha = 0{,}05$), т. е. внутри тройки различия не установлены, тогда как ROCKET и MiniRocket превосходят по этому критерию каждый метод нижней группы (разности не менее 3,19).

Суммарно по восьми наборам MiniRocket потребовалось 143,6 с, ROCKET — 736,3 с, BOSS — 41 672,7 с (11,6 ч), т. е. примерно в 5 и 290 раз больше, хотя на семи наборах из восьми MiniRocket не уступил BOSS.

Таблица 1: Accuracy на UCR и средний ранг (Major. — мажоритарный класс; при равенстве — средний ранг)

Набор	ROCKET	MiniR.	BOSS	SAX+Tr.	SAX-BoW	PatchTST	Major.
Coffee	1,000	1,000	1,000	0,893	0,643	0,488	0,536
GunPoint	1,000	0,993	1,000	0,904	0,833	0,684	0,493
Trace	1,000	1,000	1,000	0,837	0,540	0,713	0,190
TwoLeadECG	0,999	0,998	0,983	0,668	0,669	0,572	0,500
ECG5000	0,947	0,944	0,940	0,920	0,794	0,937	0,584
FordA	0,941	0,951	0,919	0,619	0,721	0,892	0,516
Wafer	0,999	0,999	0,995	0,979	0,994	0,993	0,892
ElectricDevices	0,726	0,753	0,715	0,608	0,627	0,640	0,242
Средний ранг	1,63	1,81	2,56	5,00	5,00	5,13	6,88

В опыте с прогнозированием RMSE каждой из пяти моделей оказался выше RMSE Naive во всех восьми сочетаниях актива (BTC, ETH) и горизонта H. Медиана величины $100(\mathrm{RMSE}/\mathrm{RMSE}_{\mathrm{Naive}} - 1)\,\%$ по восьми сочетаниям для отдельных моделей — от 1,27 (TiDE) до 5,19 (NHITS).

4. Заключение

На восьми выбранных наборах UCR ROCKET, MiniRocket и BOSS получили лучшие средние ранги; различия внутри этой тройки по критерию Неменьи не установлены. Суммарное время обучения и предсказания MiniRocket и BOSS в проведённых CPU-прогонах составило 2,4 мин и 11,6 ч соответственно; измерения времени носят ориентировочный характер. В отдельном опыте пять моделей прогнозирования BTC и ETH имели больший RMSE, чем Naive, во всех восьми сочетаниях актива и горизонта; статистическая значимость этих различий не оценивалась.

Для проведения исследования авторами реализованы варианты классификации SAX-BoW, SAX+Transformer и PatchTST, выполнено их сопоставление с BOSS, ROCKET и MiniRocket и проведён статистический анализ результатов. Совместный анализ качества и вычислительных затрат показал, что в проведённых экспериментах MiniRocket не уступил BOSS по accuracy на семи из восьми наборов при меньшем суммарном времени обучения и предсказания. Полученные результаты характеризуют рассматриваемые алгоритмы при использованных настройках; влияние самого представления временного ряда отдельно не выделено, а выводы не распространяются на весь архив UCR.

Список литературы

[1] P. Schäfer, “The BOSS is concerned with time series classification in the presence of noise”, *Data Mining and Knowledge Discovery*, **29**:6 (2015), 1505–1530. DOI: 10.1007/s10618-014-0377-7.

[2] A. Dempster, F. Petitjean, G. I. Webb, “ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels”, *Data Mining and Knowledge Discovery*, **34**:5 (2020), 1454–1495. DOI: 10.1007/s10618-020-00701-z.

[3] A. Dempster, D. F. Schmidt, G. I. Webb, “MiniRocket: A very fast (almost) deterministic transform for time series classification”, Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2021, 248–257. DOI: 10.1145/3447548.3467231.

[4] H. A. Dau, A. Bagnall, K. Kamgar et al., “The UCR time series archive”, *IEEE/CAA Journal of Automatica Sinica*, **6**:6 (2019), 1293–1305. DOI: 10.1109/JAS.2019.1911747.

[5] J. Demšar, “Statistical comparisons of classifiers over multiple data sets”, *Journal of Machine Learning Research*, **7** (2006), 1–30.

Статья поступила 10 октября 2026 г.

Symbolic, Numeric and Convolutional Methods for Time Series: Accuracy and Computation Time

G. V. Zhuravlev, M. I. Kumskov

Seven time-series classifiers are compared on eight UCR datasets: ROCKET, MiniRocket and BOSS form the top group by mean rank (1.63, 1.81 and 2.56) with no significant differences within it by the Nemenyi test, and the total times of MiniRocket and BOSS in the CPU runs are 2.4 min and 11.6 h, respectively. Five models forecasting BTC and ETH prices had higher RMSE than Naive in all eight asset–horizon combinations.

Keywords: time series, classification, symbolic representation, BOSS, ROCKET, MiniRocket, computational cost, forecasting.

Received on October 10, 2026