

ЧАСТОТНО-АДАПТИВНОЕ ПониЖЕНИЕ ЧАСТОТЫ ДИСКРЕТИЗАЦИИ ДАННЫХ ДЛЯ ПЕРЕНОСА МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

Алямовская А. А.^{1,2}

Научный руководитель – кандидат физико-математических наук

Мещанинов Ф.П.¹

¹АО «НИИМЭ»

²МФТИ (НИУ)

aliamovski@gmail.com

Введение

Современные модели для предсказания временных рядов хорошо зарекомендовали себя на бенчмарках с классической постановкой задачи. Однако на данный момент существует крайне актуальная область переноса моделей между доменами, сопряжённая с двумя принципиальными проблемами: распределительными сдвигами (domain shift) и различиями в частотных характеристиках сигналов (frequency shift). Классическим подходом к долгосрочному прогнозированию временных рядов является использование линейных моделей на основе декомпозиции, а также трансформеров с различными модификациями [1–4].

На практике понижение частоты дискретизации чаще всего используется как технический приём для сжатия данных и вычислительной нагрузки. При этом вопросы корректности, такие как алиасинг и сохранение спектральной структуры, как правило, игнорируются, что приводит к существенным искажениям сигнала. Опыт смежных областей показывает, что явный учёт частотной структуры перед понижением частоты дискретизации способствует повышению устойчивости и обобщающей способности нейросетей [5]. Кроме того, спектральные преобразования и выравнивание в частотной области могут рассматриваться как эффективный механизм доменной адаптации [6]. В задачах прогнозирования временных рядов также отмечается значимость совместного учёта временных и частотных признаков при доменной адаптации [7].

Основная часть

В работе разрабатывается новый метод частотно-адаптивного понижения частоты дискретизации (frequency-aware downsampling, FAD) как логичная и воспроизводимая техника повышения переносимости моделей прогнозирования между доменами с различной частотой дискретизации. В отличие от стандартных методов понижения частоты дискретизации, FAD явно учитывает спектральные свойства сигнала. Перед децимацией выполняется оценка спектра, автоматически подбирается параметр низкочастотной фильтрации на основе доли сохраняемой энергии, после чего понижение частоты осуществляется с антиалиасинговой фильтрацией. На вход FAD принимает параметры фильтрации, а также долю сохраняемой в исходном сигнале энергии. По сути, FAD выступает как модельно-независимая процедура доменной адаптации для временных рядов, устраняющая искажения, возникающие при переносе между доменами с различной частотой дискретизации, и идейно близка спектральным стратегиям доменной адаптации [6].

Для демонстрации работы разрабатываемого метода был использован широко распространённый датасет ETT [1], представляющий собой временные ряды нагрузки и

температуры двух электрических трансформаторов с различной частотой дискретизации. В качестве основного сценария рассматривается перенос моделей между доменами ETTh1 (данные с первого трансформатора) → ETTh2 (данные со второго трансформатора), где источник трансфера имеет частоту дискретизации в четыре раза выше, чем целевой домен.

Экспериментальный протокол включает предобучение моделей на домене-источнике после приведения его частоты дискретизации к целевой и оценку переносимости на целевом домене без дополнительного дообучения.

Сравнение проводилось на фиксированном наборе архитектур, охватывающем различные классы моделей: линейные базовые модели (DLinear [4]), патч-ориентированные трансформер-архитектуры (PatchTST [8]) и мультимасштабные архитектуры (TimeMixer [9]). В качестве базовых методов понижения частоты дискретизации использовались прореживание, наивное усреднение (Average Pooling) и Max Pooling. Эффективность оценивалась с использованием стандартных метрик прогнозирования (MAE и MSE) на целевом домене.

Выводы

В работе формулируется и исследуется задача переноса моделей для долгосрочного прогнозирования временных рядов: перенос между доменами с различной частотой дискретизации рассматривается как частный случай различий в частотных характеристиках сигналов. Основной вклад работы заключается в представлении FAD как простой и архитектурно-независимой процедуры доменной адаптации. Метод был протестирован на современных SOTA-архитектурах последних лет [4,8,9]. Эксперименты показывают, что использование FAD снижает ошибку переноса (среднюю разницу MSE, а также MAE) на бенчмарке ETT примерно на 3–11% по сравнению с Average Pooling и Max Pooling.

Литература

1. Zhou H. et al. Informer: Beyond efficient transformer for long sequence time-series forecasting //Proceedings of the AAAI conference on artificial intelligence. – 2021. – Т. 35. – №. 12. – С. 11106-11115.
2. Wu H. et al. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting //Advances in neural information processing systems. – 2021. – Т. 34. – С. 22419-22430.
3. Zhou T. et al. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting //International conference on machine learning. – PMLR, 2022. – С. 27268-27286.
4. Zeng A. et al. Are transformers effective for time series forecasting? //Proceedings of the AAAI conference on artificial intelligence. – 2023. – Т. 37. – №. 9. – С. 11121-11128.
5. Zhang R. Making convolutional networks shift-invariant again //International conference on machine learning. – PMLR, 2019. – С. 7324-7334.
6. Yang Y., Soatto S. Fda: Fourier domain adaptation for semantic segmentation //Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. – 2020. – С. 4085-4095.
7. He H. et al. Domain adaptation for time series under feature and label shifts //International conference on machine learning. – PMLR, 2023. – С. 12746-12774.
8. Nie Y. et al. A time series is worth 64 words: Long-term forecasting with transformers //arXiv preprint arXiv:2211.14730. – 2022.
9. Wang S. et al. Timemixer: Decomposable multiscale mixing for time series forecasting //arXiv preprint arXiv:2405.14616. – 2024.