На основе сравнения методов по MAE на наборе данных ETTh2 (720), можно выделить 4 продвинутых подходов к прогнозированию временных рядом. Перечисляю их в порядке снижения качества прогнозирования:

| Ранг | Модель | MAE | MSE | Статья | Год |
| --- | --- | --- | --- | --- | --- |
| 1 | [**SCINet** (Одномерная)](https://paperswithcode.com/paper/time-series-is-a-special-sequence-forecasting) | 0.399 | 0.249 | [Time Series is a Special Sequence: Forecasting with Sample Convolution and Interaction](https://paperswithcode.com/paper/time-series-is-a-special-sequence-forecasting) | 2021 |
| 2 | [**QuerySelector**](https://paperswithcode.com/paper/long-term-series-forecasting-with-query) | 0.4130 | 0.2585 | [Long-term series forecasting with Query Selector -- efficient model of sparse attention](https://paperswithcode.com/paper/long-term-series-forecasting-with-query) | 2021 |
| 3 | [**Informer**](https://paperswithcode.com/paper/informer-beyond-efficient-transformer-for) | 0.431 | 0.277 | [Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting](https://paperswithcode.com/paper/informer-beyond-efficient-transformer-for) | 2020 |
| 4 | [**Transformer**](https://paperswithcode.com/paper/long-term-series-forecasting-with-query) | 0.4340 | 0.2853 | [Long-term series forecasting with Query Selector -- efficient model of sparse attention](https://paperswithcode.com/paper/long-term-series-forecasting-with-query) | 2021 |
| 5 | [**SCINet** (Многомерная)](https://paperswithcode.com/paper/time-series-is-a-special-sequence-forecasting) | 0.761 | 1.074 | [Time Series is a Special Sequence: Forecasting with Sample Convolution and Interaction](https://paperswithcode.com/paper/time-series-is-a-special-sequence-forecasting) | 2021 |

1 и 5 место, отличаются тем, используется ли многомерная реализация или одномерная реализация модели. Подробнее остановимся на 4 подходах и трех статьях, где они описываются:

1. Time Series is a Special Sequence: Forecasting with Sample Convolution and Interaction[[1]](#footnote-1) (использование свёрточных нейронных сетей и сравнение их с трансформерами) Данная статья проводит дальнейшее развитие актуальных архитектур для временных рядом (рекуррентные нейронные сети, модель транформеров, сверточные нейронные сети для временных рядов). Основная идея — модифицированный подход сверточной нейронной сети, где не только расширяется зона внимания от слоя к слою, но и проводится последовательность сжатие-свертка-смешивание. Это позволяет одновременно смотреть на параметры временного ряда разного масштаба, что помогает лучше извлекать признаки из временного ряда и приводит к росту предсказательной способности в модели SCINet, предложенной авторами, в сравнении с прошлыми архитектурами.   
   На данный момент один из передовых по точности алгоритм, имеющий как многомерный так и одномерный вариант.  
   Есть официальный репозиторий[[2]](#footnote-2) с кодом и реализацией на PyTorch.
2. Long-term series forecasting with Query Selector -- efficient model of sparse attention[[3]](#footnote-3) (Использование трансформеров и query selector модели)  
   В этой статье предлагается развития идеи трансформеров за счет построения детерминистического алгоритма построения разреженной матрицы внимания.  
   Есть реализация кода от авторов модели на PyTorch[[4]](#footnote-4).
3. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting[[5]](#footnote-5) (использование информеров)  
   Данная статья разбирает проблемы прогнозирования длительных временных рядов (LSTF) и отмечает проблемы использования архитектуры трансформеров для этих целей — сложность по времени и большие требования по памяти и ограничения, присущие всем архитектурам энкодер-декодер. В качестве альтернативы, авторы предлагают более экономичную архитектуру модели для прогнозирования длительных временных рядов «Информеры», которые обладают тремя основными характеристиками 1) механизм внимания, который имеет сложность по памяти и времени; 2) механизм внимания, который выделяет ключевые признаки, одновременно с делением пополам размера входов на каскадных слоях, что позволяет эффективно работать с чрезвычайно длинными временными рядами; 3) декодер генеративного типа, который строит прогноз за один шаг, а не итеративно, чем значительно ускоряет построение прогнозов.  
   В данной статье проводится сравнения с другими методами и показано превосходство предложенного подхода в разрезе прогнозирования длительных временных рядов.  
   Есть реализация кода от авторов модели на PyTorch[[6]](#footnote-6).

Эти статьи про методы, которые не входят в state-of-the-art, но я считаю кажется, их можно упомянуть:

1. Probabilistic Forecasting with Temporal Convolutional Neural Network[[7]](#footnote-7)  
   В статье описывается подход к созданию прогнозов для мультивариативных временных рядом на основе свёрточных нейронных сетей с использованием обучения признакам.  
   Есть официальная реализация в коде от авторов работы на keras[[8]](#footnote-8), что важнее — реализован внутри пакета Darts – современный и быстроразвивающийся пакет для работы с временными рядами[[9]](#footnote-9)
2. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting[[10]](#footnote-10)   
   Основной фокус на одномерных временных рядах — минус при анализе мультивариантных временных рядов. Предложен и реализован подход на основе глубокой нейронной сети со связями вперед и назад между слоями и набором множества полносвязных слоев внутри архитектуры сети.  
   Есть реализация от компании, связанной с автором статьи[[11]](#footnote-11)   
   Также реализован внутри пакета Darts
3. Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting.[[12]](#footnote-12)  
   На этой работе в первую очередь базируется работа Liu 2021 года (модель SCINet)  
   Основная идея — использование трансформеров и обучение поведению временного ряда на разных масштабах плюс выделение отдельной части модели под интерпретируемое представление внимания обученной сети, что позволяет и достигнуть лучших результатов в нескольких бенчмарках и отойти от реализации — черный ящик.   
   Есть реализация на tensorflow[[13]](#footnote-13), статья от исследователей в google research.

1. Time Series is a Special Sequence: Forecasting with Sample Convolution and Interaction Minhao Liu, Ailing Zeng, Zhijian Xu, Qiuxia Lai, Qiang Xu https://arxiv.org/abs/2106.09305 [↑](#footnote-ref-1)
2. https://github.com/cure-lab/SCINet [↑](#footnote-ref-2)
3. Long-term series forecasting with Query Selector -- efficient model of sparse attention Jacek Klimek, Jakub Klimek, Witold Kraskiewicz, Mateusz Topolewski https://arxiv.org/abs/2107.08687v2 [↑](#footnote-ref-3)
4. https://github.com/moraieu/query-selector [↑](#footnote-ref-4)
5. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, Wancai Zhang https://arxiv.org/abs/2012.07436v3 [↑](#footnote-ref-5)
6. <https://github.com/zhouhaoyi/Informer2020> — самый популярный по активности редакторов и наблюдателей репозиторий из этого обзора [↑](#footnote-ref-6)
7. Probabilistic Forecasting with Temporal Convolutional Neural Network Yitian Chen, Yanfei Kang, Yixiong Chen, Zizhuo Wang https://arxiv.org/abs/1906.04397v3 [↑](#footnote-ref-7)
8. https://github.com/ashishpatel26/tcn-keras-Examples [↑](#footnote-ref-8)
9. https://github.com/unit8co/darts [↑](#footnote-ref-9)
10. N-BEATS: Neural basis expansion analysis for interpretable time series forecasting Boris N. Oreshkin, Dmitri Carpov, Nicolas Chapados, Yoshua Bengio [↑](#footnote-ref-10)
11. https://github.com/ElementAI/N-BEATS [↑](#footnote-ref-11)
12. Temporal Fusion Transformers for Interpretable Multi-horizon Time Series Forecasting Bryan Lim, Sercan O. Arik, Nicolas Loeff, Tomas Pfister https://arxiv.org/abs/1912.09363v3 [↑](#footnote-ref-12)
13. https://github.com/google-research/google-research/tree/master/tft [↑](#footnote-ref-13)