



ITMO UNIVERSITY

Saint Petersburg, Russia

Построение адаптивного алгоритма для прогнозирования макроэкономических данных

FEDOT team

28.08.2022

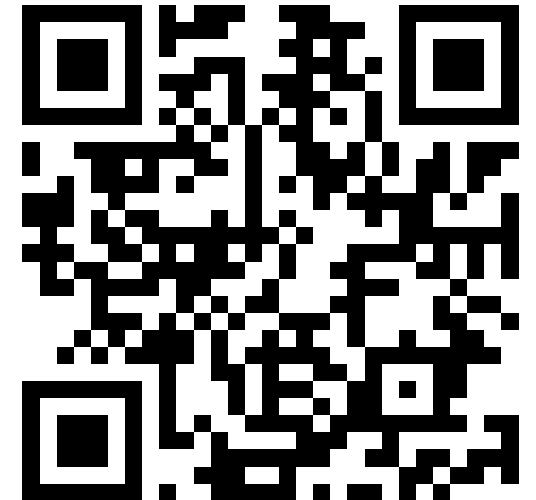
Работа, работа, иди на Федота!

Задача: научиться **быстро** и с **высокой точностью** прогнозировать макроэкономические показатели. Требуется, чтобы модель можно было быстро **адаптировать** к другим рядам

Что же делать?

Решение: композитный AutoML-фреймворк FEDOT

- Автоматический подбор структуры модели (композиция)
- Автоматическая настройка гиперпараметров (тюнинг)
- Решение задач классификации/регрессии/**прогнозирование временных рядов**
- Простой интерфейс
- Разрабатывается в ИТМО в **России**



Мета-модель

Идея: создать верхнеуровневую модель, которая сможет подбирать гиперпараметры для заданной последовательности алгоритмов (пайплайна), опираясь на уже известные ряды.

Это позволяет быстрее **адаптировать** уже обученные модели для новых рядов.

Условное предсказание

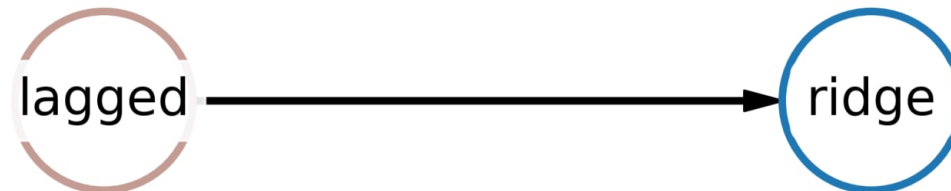
Предсказываем, учитывая известный ряд как экзогенную переменную

Эксперименты по автоматическому подбору структуры модели показали, что время, затрачиваемое на индивидуальный поиск оптимального пайплайна для каждого временного ряда, непропорционально велико

Поэтому были проверены некоторые эффективные связки моделей, известные из опыта решения подобных задач. В результате в качестве базового пайплайна используется модель авторегрессии с L2-регуляризацией по следующим причинам:

- Простота модели
- Быстродействие
- Универсальность (хорошо работает на большинстве тренировочных рядов)
- Интерпретируемость

При безусловном прогнозировании прогноз выполняется отдельно для каждого ряда, даже если они находятся в одном файле.



Наша метамодель это:

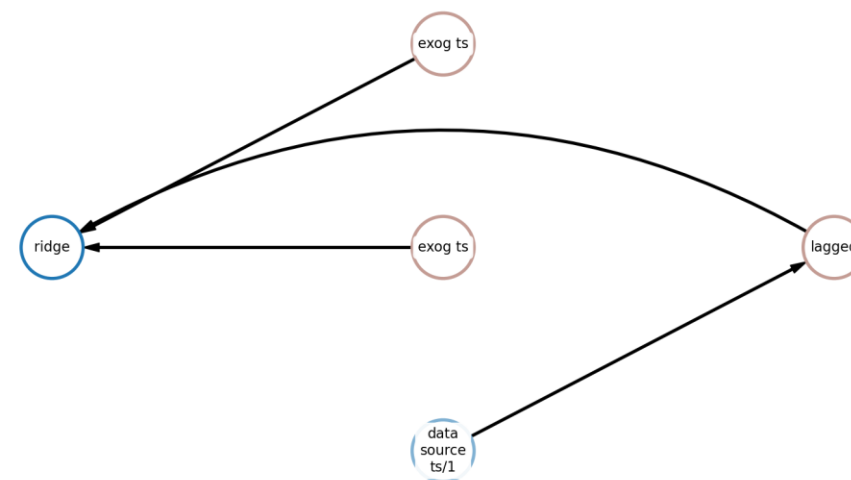
1. Для безусловного прогноза -- использование векторного представления временного ряда.

Описание: С помощью метода скользящего окна получаем значения 1, 2 и 3 квартилей и составляем из них вектор.



2. Для условного прогноза -- использование экзогенных переменных.

Описание: уже известный временной ряд берется как экзогенная переменная и используется для прогнозирования оставшихся.





Результаты. Метаобучение

MAPE, %	Var	Std	Mean	Median	Min	Max
Квантили без тюнинга	7856990	28030.3	8476.36	2996.06	98.84	235671
Корреляция без тюнинга	7856990	28030.3	8476.36	2996.06	98.84	235671
Квантили +тюнинг (10 it)	8.722	29.532	8.903	1.894	0.001	240.324
Корреляция +тюнинг (10 it)	11.612	34.077	11.788	1.772	0.009	240.463

Заметно, что подбор гиперпараметров необходим даже при использовании метаобучения, хотя метаобучение заметно ускоряет сходимость.

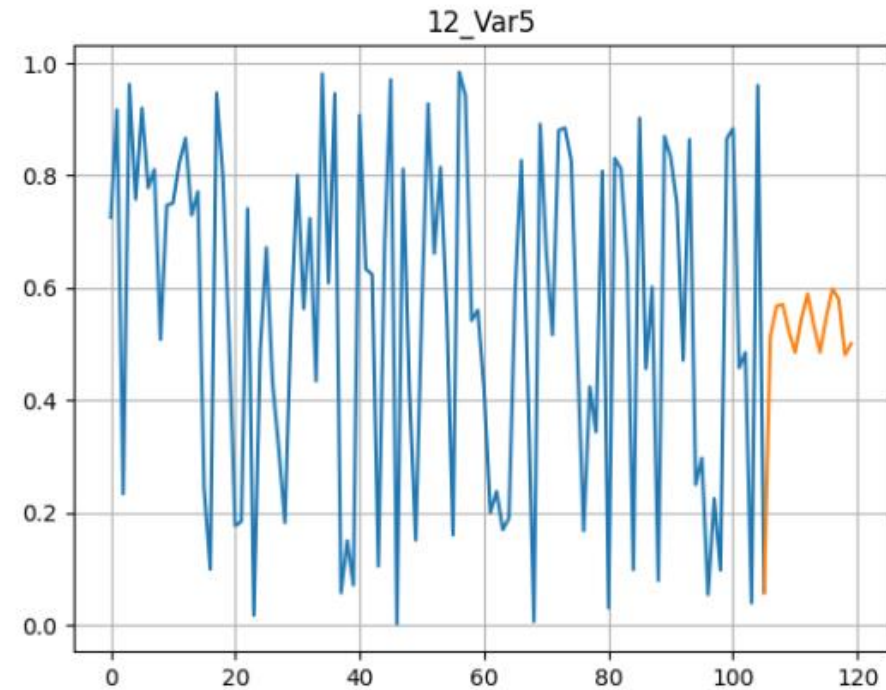
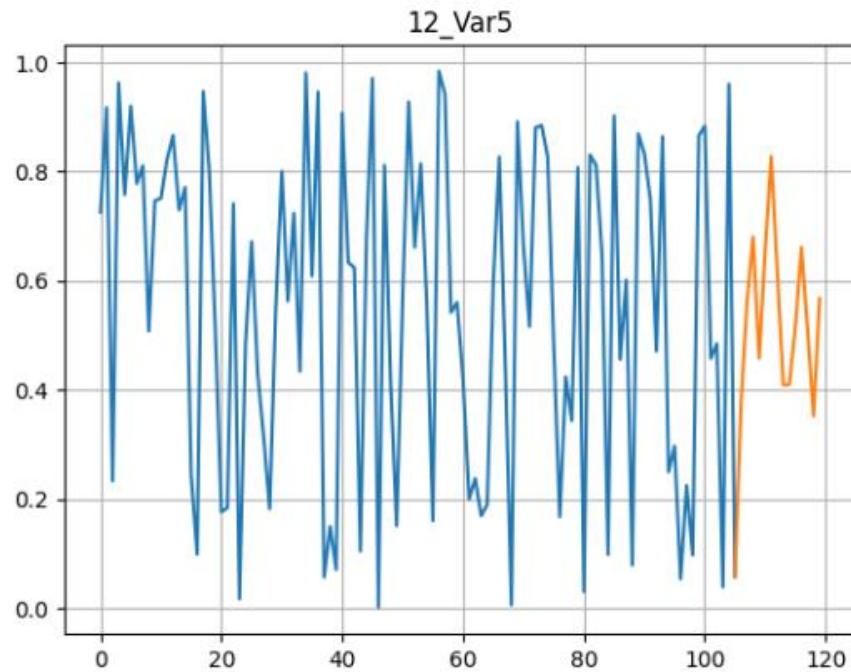
Результаты. Метаобучение

MAPE, %	Var	Std	Mean	Median	Min	Max
Без метаобучения Тюнинг (50 it)	0.027	1.648	9.738	9.738	8.090	11.387
Метаобучение+ тюнинг (10 it)	8.722	29.532	8.903	1.894	0.001	240.324
Метаобучение+ тюнинг (25 it)	8.713	29.518	8.374	1.313	0.019	241.450

Т.к. для Score более важна точность в целом, уменьшение медианы MAPE важнее, чем более точное предсказание на отдельных рядах.

Разница между 10 и 25 итерациями тюнинга имеется, но требует в 2.5 раза больше времени на обучение, что невозможно в данных временных рамках

Результаты. Экзогенная переменная



Заметно, что модель с экзогенной переменной (слева) лучше воспроизводит структуру ряда, чем модель без нее (справа)



Валерий Покровский
ИТМО, ML-инженер



Андрей Гетманов
ИТМО, ML-инженер



Майя Пинчук
ИТМО, ML-инженер

Спасибо за внимание!

IT'sMO *re than a*
UNIVERSITY