

# Применение ИИ-моделей стабильной части остатков клиентов в управлении ALM-рисками (на примере компаний Группы Московской биржи)



## **ИИ-МОДЕЛЬ ОЦЕНКИ СТАБИЛЬНОЙ ЧАСТИ ОСТАТКОВ КЛИЕНТОВ НКЦ**

**Цель внедрения ИИ-моделей:** увеличение дохода от инвестирования стабильной части остатков клиентов в облигации и МБК на сроки до 30 дней при сохранении приемлемого уровня риск-защищенности модели, т.е. соблюдения ограничений на риск-аппетит и индикативные показатели риска ликвидности и ALM-рисков, при благоприятной рыночной конъюнктуре.

**Допущения:**

- Модель применяется только к рублевым остаткам
- Сроки, для которых применима модель: 7,14 и 30 дней
- ИИ-модель применима только для сроков до 30 дней и существует параллельно исторической модели стабильной части, которая применяется на сроках свыше 30 дней и используется для построения гэта ликвидности и расчета метрик риск-аппетита по риску ликвидности

**Взаимодействие:**

1. Разработка модели – подразделение ИИ
2. Внедрение, мониторинг метрик **риск-устойчивости** и **денежные метрики**, вывод в ОПЭ и оценка влияния на РА и индикативы – подразделение рыночных рисков
3. Создание и поддержание БД с историей входных и выходных данных модели - ИТ

**Метрики риск-устойчивости (надежности)<sup>1</sup>**

*M1* = средневзвешенный прогноз стабчасти за период, руб – (“минус”) макс. пробой за этот период;  
*M2* = средневзвешенный прогноз стабчасти за период, руб – (“минус”) средний пробой за период;  
Кол-во пробоев за период;  
Частота пробоев за период (<0.1%);  
Максимальный пробой за период;  
Средний пробой за период

В дату t имел место пробой прогноза стабчасти на срок h дней, если прогноз стабчасти превысил минимальный остаток ИКО за следующие h дней. Величина пробоя в дату t = Прогноз стабчасти на h дней - минимальный остаток ИКО за следующие h дней (то есть за даты t+1, t+2, ..., t+h).

**Денежные метрики**

**M1den** показывает сальдо между гипотетическим доходом ML-модели и Hist-модели за период анализа, приведенная к одному году. Сальдо возникает от размещения дополнительных средств, спрогнозированных AI-моделью, на сроки 7/14/30 дней под более высокий процент по сравнению с размещением на 1 день:

$$M1den = \frac{1}{YF \cdot 365 \cdot 100} \times \sum_{t=Start}^{End} (\text{Прогноз\_}AI_t - \text{Прогноз\_}Hist_t) * \max(0, ROISFix\_nd_t - RUONIA_t)$$

**M2den** корректирует M1den на величину расходов, связанных с фондированием пробоев по ставке RUONIA ежедневно в течение n дней, где n = 7,14,30 дней

$$M2den = M1 - \frac{n}{YF \cdot 365 \cdot 100} \times [\sum_{t=Start}^{End} Prob\_AI_t * RUONIA_t - \sum_{t=Start}^{End} Prob\_Hist_t * RUONIA_t]$$

*Start* – дата начала периода расчета;  
*End* – дата конца периода расчета;  
*Прогноз\_AI<sub>t</sub>* – прогноз стабильной части на сроки 7/14/30 дней, выполненный ML-моделью в дату t;  
*Прогноз\_Hist<sub>t</sub>* – прогноз стабильной части на срок 7 дней, выполненный действующей Hist-моделью в дату t;  
*ROISFix\_nd<sub>t</sub>* – ставка *ROISFix* на срок n дней (в процентах, нужно делить на 100), где n = 7,14,30 дней  
*RUONIA<sub>t</sub>* – ставка *RUONIA* на 1 день (в процентах, нужно делить на 100);  
*Prob\\_AI<sub>t</sub>* – положительная разница между прогнозом ML-модели и фактическим минимальным остатком обеспечения за следующие 7/14/30 дней  
*Prob\\_Hist<sub>t</sub>* – положительная разница между прогнозом действующей Hist-модели и фактическим минимальным остатком обеспечения за следующие 7/14/30 дней;  
*YF* = (*End* – *Start*)/365 – количество лет в периоде [Start; End]

Принятие решений о применении исторической модели или ИИ-модели стабильной части остатков

\*M1, M2, M1den, M2den - в млн руб.

на 7 дней			AI-модель						Hist-модель (99,9% - персентиль)					
			Метрики надежности				Денежные метрики		Метрики надежности				Денежные метрики	
№	Начало периода	Конец периода	Кол-во пробоев	Частота пробоев,%	M1	M2	M1den	M2den	Кол-во пробоев	Частота пробоев,%	M1	M2	M1den	M2den
0	01.01.2018	22.02.2022	0	0	30 000	30 000	2,33	2,33	0	0	27 500	27 500	0	0
1	22.03.2022	24.07.2023	0	0	62 260	62 260	19,5	19,5	0	0	49 000	49 000	0	0
2	18.06.2024	08.04.2025	0	0	64 000	64 073	46	46	0	0	39 000	39 000	0	0

на 14 дней			AI-модель					
			Метрики надежности				Денежные метрики	
№	Начало периода	Конец периода	Кол-во пробоев	Частота пробоев,%	M1	M2	M1den	M2den
0	01.01.2018	22.02.2022	0	0	28 000	28 000	2	2
1	22.03.2022	24.07.2023	0	0	53 000	53 000	11	11
2	18.06.2024	08.04.2025	0	0	55 000	55 000	51	51

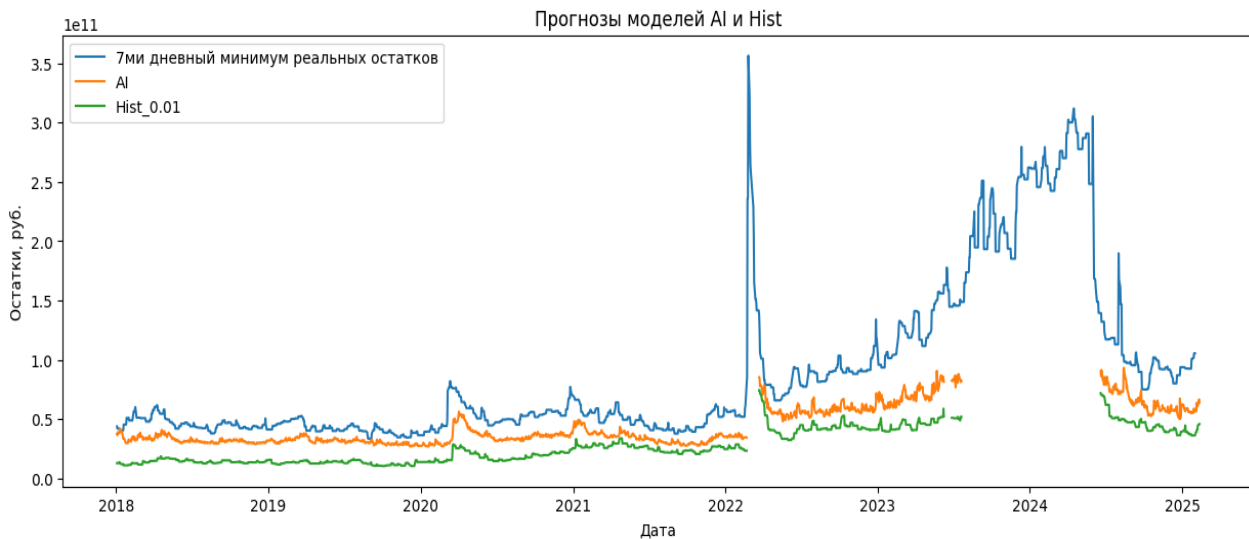
на 30 дней			AI-модель					
			Метрики надежности				Денежные метрики	
№	Начало периода	Конец периода	Кол-во пробоев	Частота пробоев,%	M1	M2	M1den	M2den
0	01.01.2018	22.02.2022	0	0	27 000	27 000	0,30	0,3
1	22.03.2022	24.07.2023	0	0	46 000	46 000	5,00	0,5
2	18.06.2024	08.04.2025	0	0	49 000	49 000	61,00	61

Обучающая выборка: с 2018 по 2023 год вкл.  
Тестовая выборка: 07.2024-12.2024  
Исключенные из обучающей выборки периоды повышенной волатильности в остатках

Качественное описание	AI-модель	Hist-модель
	Метод градиентного бустинга	Исторический метод: расчет 99%-персентилья исторического распределения разницы между минимальным остатком за следующие К дней и минимальным остатком за предыдущие 7 дней относительно даты t, которая изменяется в заданном историческом периоде с 01.01.2014 по дату расчета, где К - срок оценки стаб. части
	Крупнейшие УК из обучающей выборки не исключаются	Исключаются 2 крупнейших УК

# Графики реальных остатков и прогнозов AI и Hist на сроки 7,14 и 30 дней

на 7 дней



на 14 дней



на 30 дней



## **ИИ-МОДЕЛЬ ОЦЕНКИ СТАБИЛЬНОЙ ЧАСТИ ОСТАТКОВ ФИНУСЛУГ НРД**

# 1. Цель модели

Каждому клиенту Финуслуг на платформе **доступен кошелек**, на котором можно **хранить деньги** для открытия сберегательных продуктов и даже **зарабатывать от 17% годовых** при активации профита

**>1.3 млн.**

клиентов ФУ

**>6 млрд.**

средств на кошельках

**~57 тыс.**

клиентов ФУ активировали  
профит

**~78 тыс.**

клиентов ФУ хранят деньги на  
кошельке

## Задача:

Прогнозировать объем денежных средств на кошельках клиентов ФУ на период 7, 14 и 30 дней для инвестирования средств на МБК (депозиты) с целью получения дополнительного дохода.

## Задача машинного обучения:

Разработать модель прогнозирования стабильной части остатков средств на кошельках с минимальной процентной ошибкой (MAPE) и абсолютной ошибкой (MAE).

## 2. Алгоритм разработки модели

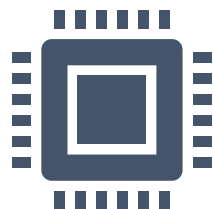
### Шаг 1:



**Сбор данных**, которые могут объяснить актуальный объем денежных средств на кошельках.

>100 уникальных факторов, предоставляемых подразделением рыночных рисков

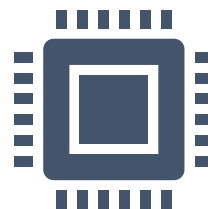
### Шаг 2:



**Разработка ML модели**, которая наилучшим образом “подстроится” под данные.

~60 лучших параметров отобрано для обучения модели

### Шаг 3:



**Тестирование ML модели** на исторических данных, оценка качества работы модели.

24 прогноза по 7 дней формировали метрики

### Шаг 4:



**Прогноз стабильной части на 7, 14 и 30 дней**

+ проведение бэктеста модели для расчета “пробоев”

расчет X руб., которые можно инвестировать



### 3.Сбор данных факторов модели: поведенческие макроэкономические и бизнес-факторы



Используемые параметры

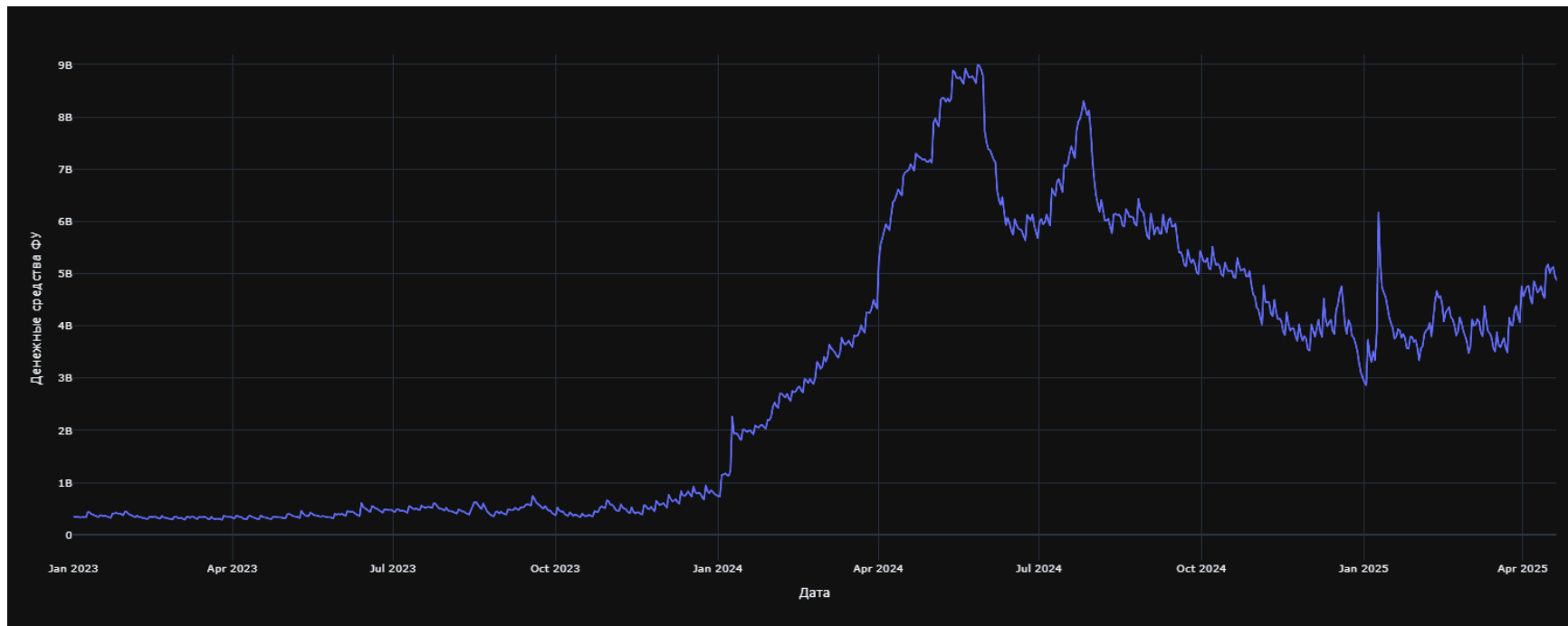
- 58** параметров – транзакции
- 12** параметров – календарь
- 12** параметров – макрофакторы
- 9** параметров – ключевые ставки
- 7** параметров – профит
- 6** параметров – депозит
- 4** параметра – кошелек
- 1** параметр – целевая переменная

С учетом лагов и скользящих средних  
всего

**764** параметра

## 4.Целевая переменная (NEW)

График объема денежных средств ФУ с течением времени



Статистические  
свойства

Max = 12,9 млрд. руб.  
Mean = 2.66 млрд. руб.  
Std = 3.07 млрд. руб.

Тест Дики-Фуллера\*

ADF = 0.38  
Critical Value (5%) = -2.864  
P-value = 0.980 (>0.05)

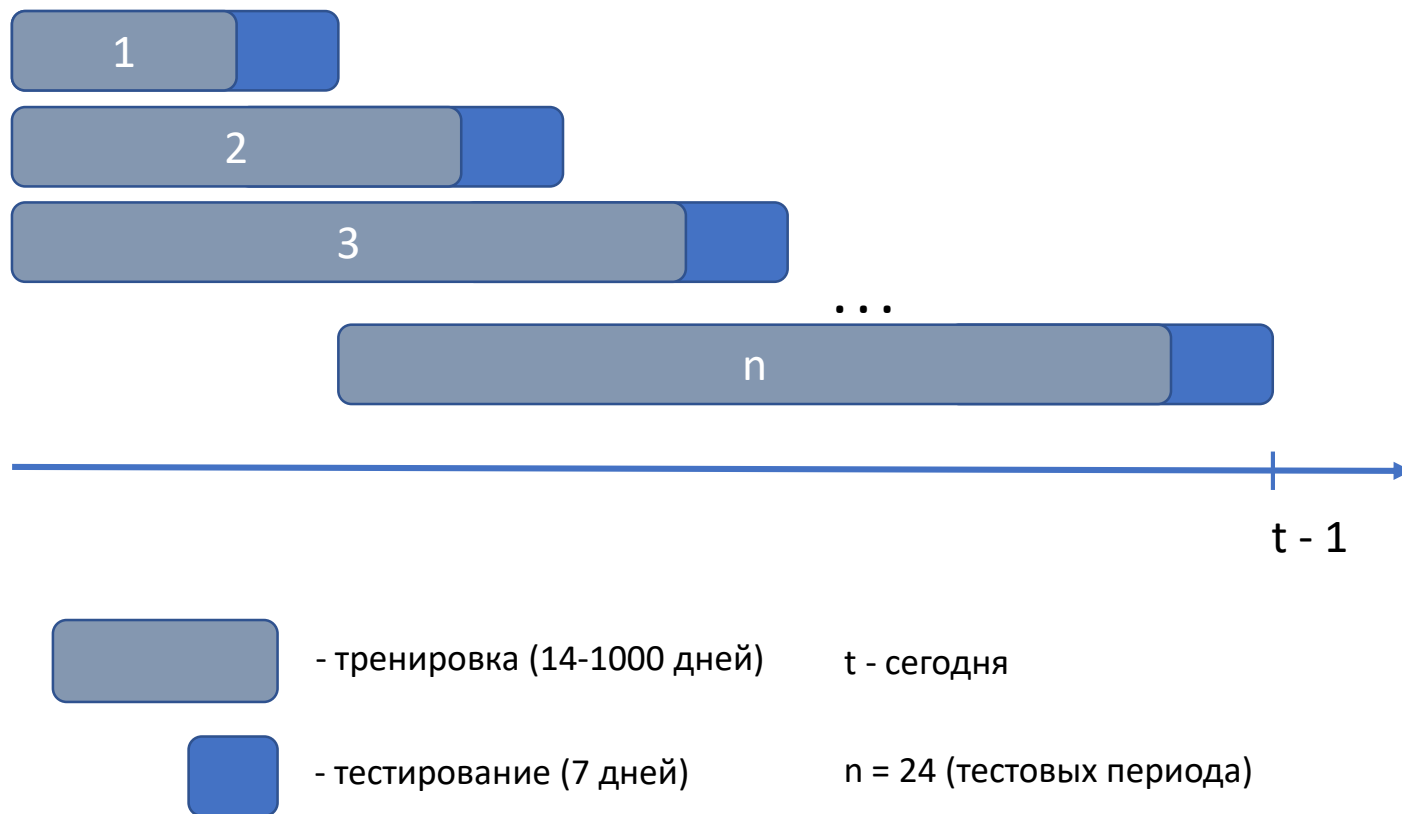
Выводы о целевой  
переменной

Статистические свойства  
целевой переменной  
меняются со временем

# 5.1 Разработка модели (NEW)

методика тестирования модели  
на исторических данных

## Кросс-валидация (скользящее окно)



### Шаг 1

По каждому из  $n$  тестовых периодов  
рассчитать метрики:

1.1 **MAPE** – средняя абсолютная  
процентная ошибка

1.2 **MAE** – средняя процентная  
ошибка

1.3 **OV** (over-predictions) –  
количество пробоев

### Шаг 2

Фиксируем все показатели и  
усредняем

### Шаг 3

Получаем средние значения метрик  
на всем тестовом периоде на  
основании исторических данных

## 5.2 Разработка модели

выбор модели и настройка параметров

### ML-модель

**Catboost** подходит для прогнозирования временных рядов благодаря способности обрабатывать большие объёмы данных со сложными закономерностями, улавливать краткосрочные и долгосрочные зависимости.

Выбираем модель  
Catboost Regressor

+

### Лучшие гиперпараметры

#### Optuna

это платформа оптимизации гиперпараметров, которая позволяет эффективно настраивать модели машинного обучения. Используя optuna, тестируем комбинацию гиперпараметров по методике кросс-валидации.

Минимизируя MAPE, выбираем лучшую комбинацию гиперпараметров

+

### Лучшие параметры (фичи)

#### Feature Importance

это оценка, которая присваивается каждому признаку и указывает на его относительный вклад в прогнозы.

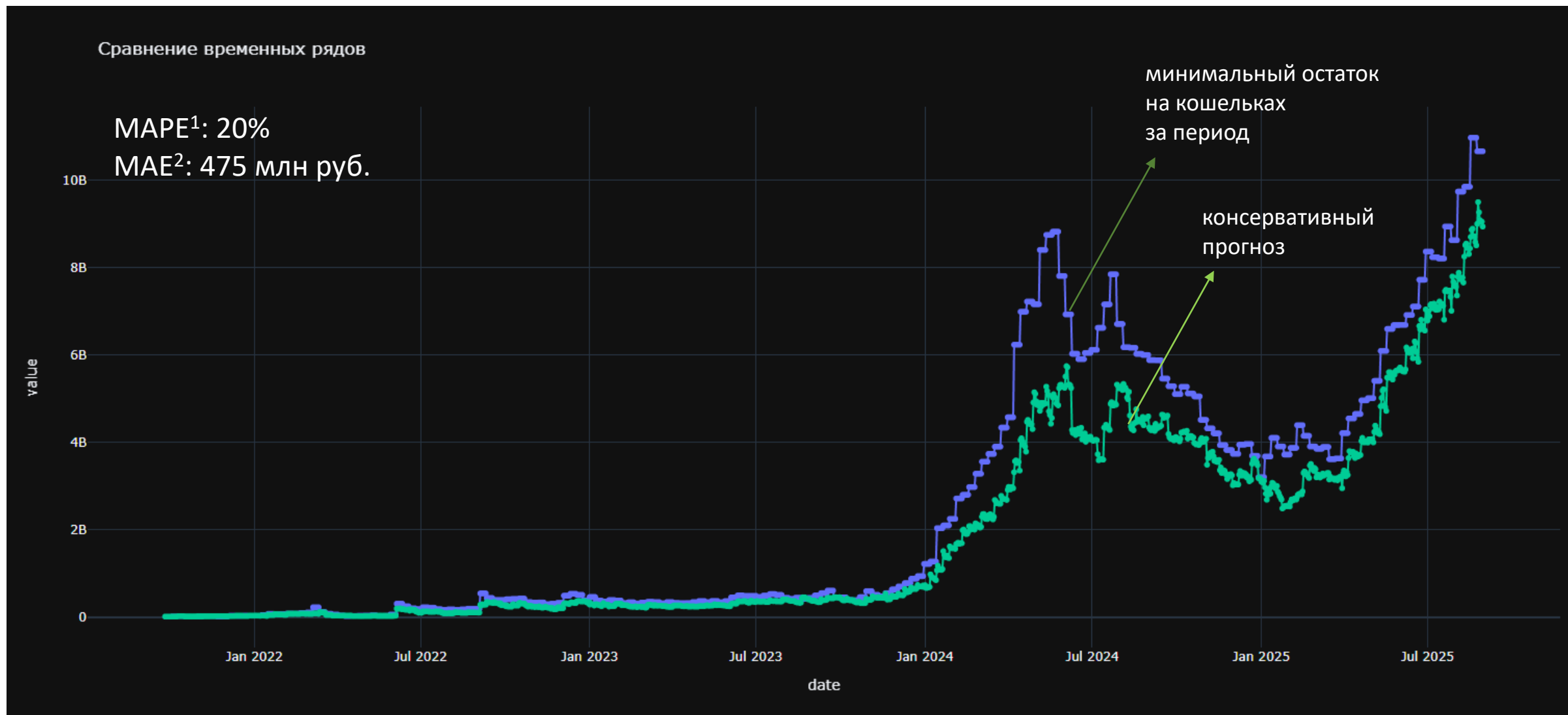
По методике кросс-валидации на каждый тестовый период рассчитываются свои топ-100 параметров.

Параметры, которые попадали минимум 8 раз в списки топ-100, выбираем в итоговый список лучших параметров

=

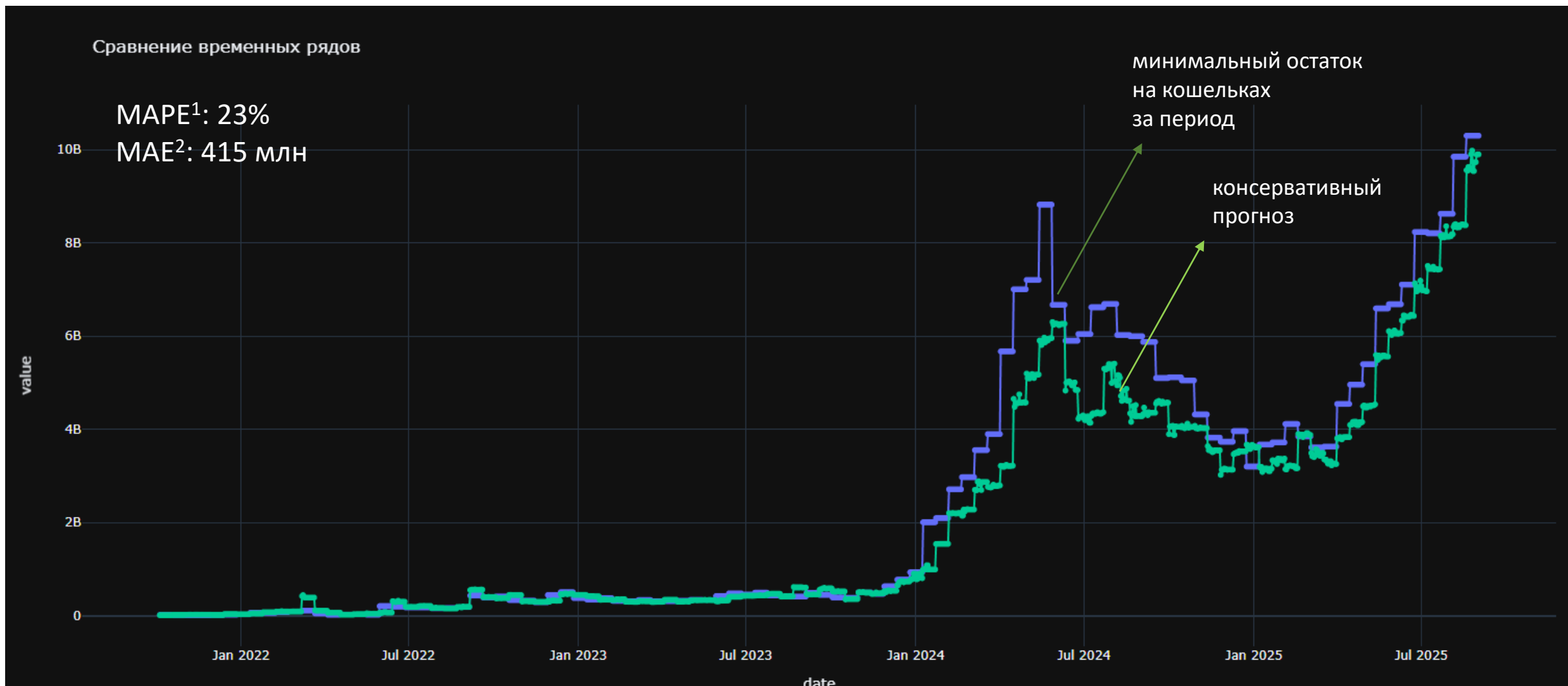
**ML-модель,**  
готовая к тестированию  
на исторических  
данных

## 5.6 Разработка модели (консервативный прогноз 7 дней)



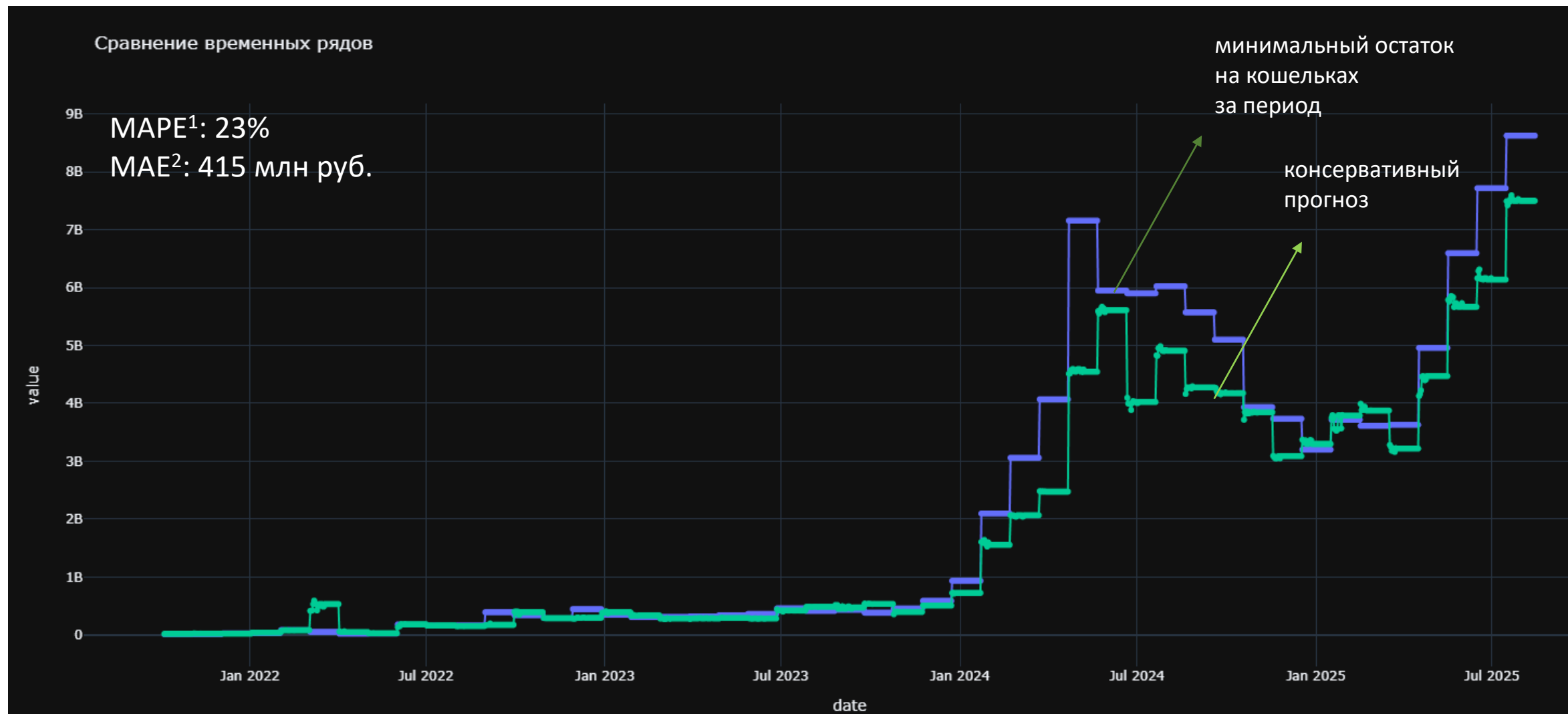
1) MAPE – средняя абсолютная ошибка (%) 2) MAE – средняя абсолютная ошибка

## 5.7 Разработка модели (консервативный прогноз 14 дней)



1) MAPE – средняя абсолютная ошибка (%) 2) MAE – средняя абсолютная ошибка

## 5.8 Разработка модели (консервативный прогноз 30 дней)



1) MAPE – средняя абсолютная ошибка (%) 2) MAE – средняя абсолютная ошибка