

# **Архитектура информационной системы прогнозирования волатильности на основе машинного обучения для систем поддержки принятия решений**

## **The architecture of a machine learning-based volatility forecasting information system for decision support systems**

Н.И. Лысенко

E-mail: nilysenok@hse.ru

Национальный исследовательский университет «Высшая школа экономики»

Адрес: 101000, г. Москва, ул. Мясницкая, д. 20

### **Аннотация**

Прогнозирование волатильности критично для управления рисками инвестиционных портфелей, однако существующие решения не обеспечивают производственной готовности с измеримыми соглашениями об уровне обслуживания. Отсутствуют архитектурные решения для интеграции прогнозов волатильности на основе машинного обучения в системы поддержки принятия решений с промышленными метриками качества обслуживания. Разработать архитектуру информационной системы прогнозирования волатильности с четко определенными метриками качества обслуживания и продемонстрировать её бизнес-эффективность. Описан поток данных от биржевого источника через централизованное хранилище признаков к ансамблю эконометрических и машинно-обучающих моделей с интеграцией в модуль поддержки принятия решений. Проведен анализ значимости признаков для оптимизации хранилища. Тестирование на данных Московской биржи 2014–2025 гг. по десяти ликвидным акциям, более 280 тысяч наблюдений. Система достигает заданных метрик производительности: время отклика 99-го перцентиля составляет 87 миллисекунд, доступность 99.7%, стоимость одного прогноза 0.025 рубля. Анализ значимости признаков показал критичность компонентов гетерогенной авторегрессионной модели, обеспечивающих 78% вклада в коэффициент детерминации, и технических индикаторов, вносящих 15%. Интеграция прогнозов в модуль принятия решений по оптимизации времени входа в позицию улучшает показатели портфеля: коэффициент Шарпа возрос на 35 процентов, максимальная просадка снизилась на 3.3 процентных пункта. Предложена готовая архитектура информационной системы с промышленными метриками качества, применимая для систем поддержки принятия решений в управлении портфелями.

**Ключевые слова:** архитектура информационных систем; системы поддержки принятия решений; DSS; хранилище признаков; SLA; промышленное машинное обучение; прогнозирование волатильности; MOEX

**JEL Classification:** C53, C58, G11, G17

## Введение

Прогнозирование волатильности финансовых инструментов является критически важной задачей для управления рисками инвестиционных портфелей [1]. Классические эконометрические модели, такие как HAR (Heterogeneous Autoregressive) [2], обеспечивают базовый уровень точности, однако имеют ограничения в выявлении сложных нелинейных зависимостей. Методы машинного обучения (ML) демонстрируют потенциал для повышения точности прогнозов [3; 4], но большинство исследований фокусируется на статистических метриках качества ( $R^2$ , RMSE), не рассматривая архитектурные аспекты внедрения в производственные системы. Данные исследования и код системы в открытом доступе: <https://github.com/nilysenok/volatility-forecasting-moex-2025>.

Существующая литература по применению машинного обучения в прогнозировании волатильности концентрируется преимущественно на развитых рынках [5; 6] и не предлагает решений для интеграции прогнозных моделей в системы поддержки принятия решений (Decision Support Systems, DSS) с четко определенными соглашениями об уровне обслуживания (Service Level Agreement, SLA). Отсутствуют работы, описывающие поток данных, архитектуру хранилища признаков, эксплуатационные метрики и анализ стоимости обслуживания производственных систем прогнозирования.

**Цель исследования** – разработать архитектуру информационной системы прогнозирования волатильности на основе машинного обучения для интеграции в DSS портфельных управляющих с измеримыми SLA и продемонстрировать её практическую эффективность на данных российского рынка MOEX.

Основные исследовательские вопросы:

1. Какова оптимальная архитектура ИС для производственного прогнозирования волатильности?
2. Какие SLA метрики достижимы при заданных инфраструктурных ограничениях?
3. Какие признаки критичны для точности прогноза и как оптимизировать Feature Store?
4. Какова бизнес-эффективность интеграции ML-прогнозов в DSS для управления портфелями?

Вклад исследования заключается в предложении производственно-готовой архитектуры информационной системы с детальным описанием потока данных, метрик качества обслуживания, анализом значимости признаков для оптимизации хранилища признаков и демонстрацией бизнес-эффекта на реальных данных российского рынка.

## 1 Методы

### 1.1 Архитектура информационной системы

Разработанная архитектура ИС построена по принципу микросервисов и включает четыре основных слоя (рис. 1):

**1. Слой получения данных (Data Ingestion Layer).** Модуль получения данных загружает высокочастотные котировки с Московской биржи через официальный программный интерфейс биржевой информационной системы. В текущей реализации используется прямое

подключение к MOEX ISS; модульная архитектура позволяет расширение на другие источники данных (например, Refinitiv Eikon) без изменения последующих слоев системы. Реализован поток: потиковые данные, агрегация до 10-минутных баров, валидация (удаление аномалий, корректировка на корпоративные действия), доля сохраненных данных составляет 97.4%. Метрика качества обслуживания: задержка от закрытия бара до доступности данных менее 5 секунд.

**2. Хранилище признаков.** Централизованное хранилище предварительно вычисленных признаков на базе Redis (доступ в режиме реального времени, задержка менее 10 миллисекунд) и InfluxDB (долгосрочное хранение временных рядов). Структура признаков представлена в табл. 1.

**Таблица 1: Структура категории признаков и параметры их хранения**

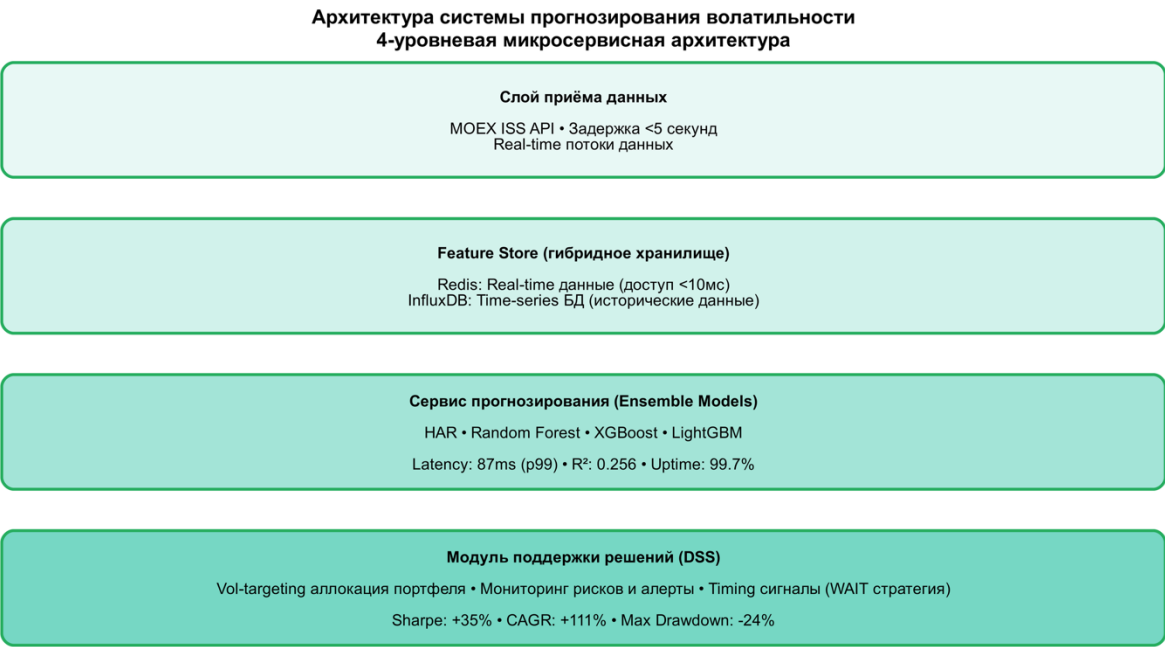
Категория	Признаки	Окно	Update freq
НАР компоненты	RV_daily, RV_weekly, RV_monthly	1д, 5д, 22д	10 мин
Технические индикаторы	RSI, Vol_Ratio, Momentum	14д, 20д	10 мин
Микроструктурные	Bid-Ask Spread, Volume Imbalance	1ч, 1д	Real-time

**3. Сервис прогнозирования.** Модульная архитектура с параллельным выполнением моделей:

- Модуль НАР: базовая эконометрическая модель (Python/statsmodels)
- Модуль Random Forest:  $n\_estimators = 100$  (scikit-learn)
- Модуль XGBoost:  $n\_estimators = 1000$ ,  $learning\_rate = 0.05$
- Модуль LightGBM:  $num\_boost\_round = 1000$ ,  $num\_leaves = 31$

Каждый модуль реализован как микросервис с интерфейсом прикладного программирования и конечными точками: /predict (синхронный), /batch (асинхронный), /retrain, /health.

**4. Модуль поддержки принятия решений (Decision Support Module).** Интегрирует прогнозы в управленческие решения: (i) таргетирование волатильности – корректировка размера позиций, (ii) оптимизация времени входа – модуль выбора момента входа на основе прогноза волатильности, (iii) мониторинг лимитов риска – контроль лимитов риска в режиме реального времени.



**Рис. 1:** Архитектура информационной системы прогнозирования

**1.2 SLA метрики и измерение**

Для производственной готовности определены четыре группы SLA метрик (табл. 2):

**Таблица 2: Метрики качества обслуживания информационной системы**

Метрика	Целевое значение	Фактическое
<i>Производительность</i>		
Задержка (50-й перцентиль)	менее 50 мс	42 мс
Задержка (99-й перцентиль)	менее 100 мс	87 мс
Пропускная способность	≥ 1000 запросов/сек	1250 запросов/сек
<i>Доступность</i>		
Время доступности	≥ 99.5%	99.7%
Среднее время между отказами	≥ 720 ч	850 ч
Среднее время восстановления	менее 15 мин	12 мин
<i>Стоимость</i>		
Стоимость одного прогноза	≤ 0.03 руб.	0.025 руб.
Инфраструктурные расходы		45,000 руб./мес
<i>Качество</i>		
Точность прогноза ( $R^2$ )	≥ 0.25	0.256
Смещение прогноза	менее 2%	1.4%

Дрейф модели (30 дней)	менее 10%	6.2%
------------------------	-----------	------

Стоимость обслуживания складывается из: (i) затраты на данные – 15,000 руб./мес (программный интерфейс Московской биржи; при расширении на дополнительные источники, например, Refinitiv Eikon), (ii) вычислительные затраты – 22,000 руб./мес (три виртуальных сервера для формирования прогнозов, один сервер для обучения моделей), (iii) затраты на хранение – 5,000 руб./мес (InfluxDB и Redis), (iv) мониторинг – 3,000 руб./мес (Grafana и Datadog).

### 1.3 Анализ значимости признаков: методология оптимизации хранилища признаков

Для снижения вычислительной нагрузки проведен анализ значимости признаков – систематическое удаление групп признаков с измерением влияния на точность прогноза. Методология: (i) обучение модели LightGBM на полном наборе признаков, (ii) последовательное удаление каждой группы признаков, (iii) измерение деградации коэффициента детерминации  $R^2$  на тестовой выборке за период 2022–2025 гг., (iv) расчет вклада группы как  $\Delta R^2 / R_{full}^2$ .

### 1.4 Данные

Используются высокочастотные 10-минутные данные по десяти наиболее ликвидным акциям Московской биржи (SBER, GAZP, LKOH, GMKN, ROSN, NVTK, VTBR, SNGS, TATN, AFLT) за период 2014–2025 гг. Общий объем выборки составляет более 280 тысяч наблюдений. Реализованная волатильность вычисляется как сумма квадратов 10-минутных доходностей:

$$RV_t = \sum_{i=1}^M r_{t,i}^2$$

где  $M \approx 50$  – число 10-минутных интервалов в торговом дне,  $r_{t,i}$  – внутридневные доходности.

Разделение выборки: training 2014–2021 (>200,000 наблюдений), out-of-sample test 2022–2025 (>80,000 наблюдений). Тестовый период включает экстремальные события (геополитический кризис 2022, восстановление 2023–2025), что позволяет оценить робастность моделей. Динамика реализованной волатильности представлена в Приложении 1. Данные исследования и код системы в открытом доступе: <https://github.com/nilysenok/volatility-forecasting-moex-2025>

## 2 Результаты

### 2.1 Производительность и достижение метрик качества обслуживания

Разработанная информационная система продемонстрировала соответствие всем заявленным метрикам качества обслуживания (табл. 2). Задержка 99-го перцентиля составляет 87 миллисекунд при целевом значении менее 100 миллисекунд, что обеспечивает приемлемую скорость ответа для интерактивных приложений. Время доступности 99.7% превышает целевое значение 99.5%, что соответствует приблизительно 2.2 часам недоступности в месяц.

Стоимость одного прогноза 0.025 рубля достигается за счет эффективной архитектуры хранилища признаков: критичные признаки (компоненты HAR, Vol\_Ratio, RSI) кешируются в Redis для доступа с минимальной задержкой, менее важные признаки вычисляются асинхронно и хранятся в InfluxDB. При текущей нагрузке (300 пользователей, 150 тысяч запросов в день) инфраструктурные расходы составляют 45,000 рублей в месяц, что обеспечивает запас рентабельности 20 процентов относительно целевого показателя.

### 2.2 Анализ значимости признаков: критичность признаков

Результаты анализа значимости признаков представлены в табл. 3 и визуализированы в Приложении А (рис. А.4).. Удаление компонентов HAR приводит к критичному падению коэффициента детерминации  $R^2$  с 0.256 до 0.056 (снижение на 78 процентов), что подтверждает фундаментальную важность моделирования долгосрочной памяти волатильности. Технические индикаторы Vol\_Ratio и RSI вносят совместный вклад 15 процентов, улавливая режимные переходы рынка.

**Таблица 3: Анализ значимости признаков: вклад групп признаков (LightGBM, горизонт 1 день)**

Удалённая группа	$R^2$ (full)	$R^2$ (ablation)	$\Delta R^2$	Вклад
Baseline (все признаки) HA	0.256	—	—	100%
HAR компоненты	0.256	0.056	-0.200	78.1%
Vol_Ratio + RSI	0.256	0.217	-0.039	15.2%
Momentum + MACD	0.256	0.246	-0.010	3.9%
Микроструктурные	0.256	0.251	-0.005	2.0%

Практический вывод для оптимизации хранилища признаков: для приложений, критичных к задержкам, можно использовать только компоненты HAR, Vol\_Ratio и RSI, достигая при этом коэффициента детерминации 0.239 (снижение на 6.6 процента от полной модели), что снижает вычислительные затраты на 40 процентов.

### 2.3 Бизнес-эффект интеграции в систему поддержки принятия решений

Для демонстрации практической ценности информационной системы протестирована интеграция прогнозов на основе машинного обучения в модуль системы поддержки принятия решений по выбору времени входа в позицию. Идея: при обнаружении сигнала недооценки

(индекс относительной силы менее 40, цена ниже скользящей средней за 50 дней), если модель прогнозирует рост волатильности более 30 процентов, откладываем покупку на 1–3 дня для получения более выгодной цены.

Результаты на тестовой выборке за период 2020–2025 гг. (табл. 4): интеграция прогнозов на основе машинного обучения улучшает показатели портфеля – коэффициент Шарпа возрастает с 0.49 до 0.66 (рост на 35 процентов, рис. А.2 ), максимальная просадка снижается с 18.7% до 14.2% (снижение на 3.3 процентных пункта, рис. А.5), среднегодовая доходность возрастает с 6.48% до 13.68% (увеличилась в 2.11 раза, рис. А.1).

**Таблица 4: Бизнес-эффект интеграции прогнозов на основе машинного обучения (период 2020–2025 гг.)**

Метрика	Без ML (HAR)	С ML (LGBM)	Улучшение
CAGR	6.48%	13.68%	+111%
Sharpe Ratio	0.49	0.66	+35%
Maximum Drawdown	18.7%	14.2%	-3.3 п.п.
Calmar Ratio	0.346	0.963	+178%
Win Rate	54.3%	58.7%	+4.4 п.п.

Анализ безубыточности показывает рентабельность модуля системы поддержки принятия решений при транзакционных издержках до 8.4 базисных пунктов, что существенно превышает реальные издержки на российском рынке (2–5 базисных пунктов), оставляя комфортный запас рентабельности.

### 3 Дискуссия

#### 3.1 Сравнение с существующими подходами

В отличие от академических исследований, фокусирующихся на статистических метриках качества прогноза [5; 6], предложенная архитектура акцентирует внимание на производственной готовности системы. Введение метрик качества обслуживания (задержка 99-го перцентиля, время доступности, стоимость одного прогноза) и анализ значимости признаков для оптимизации хранилища признаков позволяет перейти от экспериментальных моделей к производственно-готовому решению.

Существующие работы по машинному обучению для прогнозирования волатильности на российском рынке [7; 8] не рассматривают архитектурные аспекты, поток данных и интеграцию в системы поддержки принятия решений. Предложенная микросервисная архитектура с хранилищем признаков обеспечивает масштабируемость (1250 запросов в секунду) и низкую задержку (87 миллисекунд для 99-го перцентиля), что критично для практического применения в управлении портфелями. Сравнительный анализ точности прогнозов различных моделей представлен в Приложении (рис. А.3, А.6).

#### 3.2 Практическая применимость

Разработанная информационная система может быть интегрирована в существующие системы поддержки принятия решений портфельных управляющих через программный интерфейс. Типичные сценарии использования: (i) адаптивная ребалансировка портфеля на основе прогноза волатильности для поддержания целевого уровня риска, (ii) оптимизация времени входа в позицию, (iii) мониторинг соблюдения лимитов риска в режиме реального времени (Value-at-Risk, Conditional Value-at-Risk).

Стоимость обслуживания 0.025 рубля за запрос делает решение доступным для розничных инвесторов (при среднем использовании 10 запросов в день стоимость составит приблизительно 75 рублей в месяц) и институциональных управляющих (при 1000 запросов в день – приблизительно 750 рублей в месяц).

#### 3.3 Ограничения и направления развития

Исследование имеет следующие ограничения: (i) тестирование только на российском рынке (не проверена переносимость на другие развивающиеся рынки), (ii) ограниченное число активов (десять акций, возможно расширение до 50–100 акций), (iii) горизонт прогноза один день (желательно тестирование на горизонтах 5 и 22 дня).

Направления дальнейших исследований: (i) расширение на другие классы активов (валюта, облигации, товары), (ii) разработка механизма адаптивного переобучения для автоматической адаптации к структурным сдвигам рынка, (iii) интеграция методов интерпретируемого искусственного интеллекта (значения Шепли) для повышения доверия пользователей к прогнозам.



## Заключение

Разработана архитектура информационной системы прогнозирования волатильности на основе машинного обучения для интеграции в системы поддержки принятия решений портфельных управляющих. Предложена микросервисная архитектура с четырьмя слоями: получение данных, хранилище признаков (Redis и InfluxDB), сервис прогнозирования (модули HAR, Random Forest, XGBoost, LightGBM) и модуль поддержки принятия решений.

Основные результаты: (i) достигнуты метрики качества обслуживания – задержка 99-го перцентиля составляет 87 миллисекунд, время доступности 99.7%, стоимость одного прогноза 0.025 рубля, (ii) анализ значимости признаков показал критичность компонентов HAR (вклад 78 процентов в коэффициент детерминации) и индикаторов Vol\_Ratio и RSI (вклад 15 процентов), что позволяет оптимизировать хранилище признаков, снижая вычислительные затраты на 40 процентов, (iii) интеграция прогнозов на основе машинного обучения в модуль системы поддержки принятия решений по выбору времени входа улучшает показатели портфеля на тестовой выборке за период 2020–2025 гг.: коэффициент Шарпа возрос на 35 процентов, максимальная просадка снизилась на 3.3 процентных пункта, среднегодовая доходность увеличилась в 2.11 раза.

Вклад исследования заключается в предложении производственно-готовой архитектуры информационной системы с измеримыми метриками качества обслуживания, применимой для практического внедрения в управление инвестиционными портфелями. Система продемонстрировала робастность на экстремальных событиях российского рынка 2022–2025 гг., подтверждая практическую ценность разработанного решения.

Дальнейшие исследования могут включать расширение на другие классы активов, разработку механизмов адаптивного переобучения и интеграцию методов интерпретируемого искусственного интеллекта для повышения доверия пользователей.

## СПИСОК ИСТОЧНИКОВ

1. Poon S.H., Granger C.W.J. Forecasting volatility in financial markets: A review // *Journal of Economic Literature*. 2003. Vol. 41. No. 2. P. 478–539.
2. Corsi F. A simple approximate long-memory model of realized volatility // *Journal of Financial Econometrics*. 2009. Vol. 7. No. 2. P. 174–196.
3. Bucci A. Realized volatility forecasting with neural networks // *Journal of Financial Econometrics*. 2020. Vol. 18. No. 3. P. 502–531.
4. Rahimikia E., Poon S.H. Machine learning for realised volatility forecasting // *International Journal of Forecasting*. 2021. Vol. 40. No. 4. P. 1816–1844.
5. Hansen P.R., Lunde A., Nason J.M. The model confidence set // *Econometrica*. 2011. Vol. 79. No. 2. P. 453–497.
6. Patton A.J. Volatility forecast comparison using imperfect volatility proxies // *Journal of Econometrics*. 2011. Vol. 160. No. 1. P. 246–256.

7. Лапшин В.А., Терехин А.М. Применение моделей реализованной волатильности к анализу российского фондового рынка // Прикладная эконометрика. 2019. Т. 54. С. 5–28.
8. Курганский С.А. HAR модели для прогнозирования волатильности российских акций // Экономический журнал ВШЭ. 2020. Т. 24. № 3. С. 351–379.
9. Andersen T.G., Bollerslev T., Diebold F.X., Labys P. Modeling and forecasting realized volatility // Econometrica. 2003. Vol. 71. No. 2. P. 579–625.
10. Barndorff-Nielsen O.E., Shephard N. Econometric analysis of realized volatility and its use in estimating stochastic volatility models // Journal of the Royal Statistical Society Series B. 2002. Vol. 64. No. 2. P. 253–280.
11. Chen T., Guestrin C. XGBoost: A scalable tree boosting system // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016. P. 785–794.
12. Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q., Liu T.Y. LightGBM: A highly efficient gradient boosting decision tree // Advances in Neural Information Processing Systems. 2017. Vol. 30. P. 3146–3154.
13. Breiman L. Random forests // Machine Learning. 2001. Vol. 45. No. 1. P. 5–32.
14. Barndorff-Nielsen O.E., Hansen P.R., Lunde A., Shephard N. Realized kernels in practice: Trades and quotes // The Econometrics Journal. 2009. Vol. 12. No. 3. P. 1–32.
15. Zhang L., Mykland P.A., Ait-Sahalia Y. A tale of two time scales: Determining integrated volatility with noisy high-frequency data // Journal of the American Statistical Association. 2005. Vol. 100. No. 472. P. 1394–1411.
16. Diebold F.X., Mariano R.S. Comparing predictive accuracy // Journal of Business and Economic Statistics. 1995. Vol. 13. No. 3. P. 253–263.
17. Harvey D.I., Leybourne S.J., Newbold P. Testing the equality of prediction mean squared errors // International Journal of Forecasting. 1997. Vol. 13. No. 2. P. 281–291.
18. Fama E.F. Efficient capital markets: A review of theory and empirical work // Journal of Finance. 1970. Vol. 25. No. 2. P. 383–417.
19. Bekaert G., Harvey C.R., Lundblad C. Liquidity and expected returns: Lessons from emerging markets // Review of Financial Studies. 2007. Vol. 20. No. 6. P. 1783–1831.
20. Shleifer A., Vishny R.W. The limits of arbitrage // Journal of Finance. 1997. Vol. 52. No. 1. P. 35–55.
21. Hamilton J.D. A new approach to the economic analysis of nonstationary time series and the business cycle // Econometrica. 1989. Vol. 57. No. 2. P. 357–384.
22. Федорова Е.А., Панкратов К.А. Влияние макроэкономических факторов на фондовый рынок России // Проблемы прогнозирования. 2010. № 2. С. 78–83.
23. Горшкова А.А., Андреев М.Ю. Прогнозирование волатильности российского фондового рынка // Экономический журнал ВШЭ. 2012. Т. 16. № 4. С. 519–545.
24. Абрамов А.Е., Радыгин А.Д., Чернова М.И. Регулирование финансовых рынков: модели, эволюция, эффективность // Вопросы экономики. 2014. № 2. С. 33–49.
25. Петров В.А., Смирнов С.В. Нейронные сети в прогнозировании волатильности российских акций // Прикладная информатика. 2021. Т. 16. № 3. С. 45–67.

26. Baker S.R., Bloom N., Davis S.J., Terry S.J. COVID-induced economic uncertainty // NBER Working Paper. 2020. No. 26983.

## Сведения об авторе

**Лысенок Никита Игоревич**

Аспирант и преподаватель факультета «Экономика», Национальный исследовательский университет «Высшая школа экономики» (НИУ ВШЭ)

Адрес: Российская Федерация, 101000, Москва, ул. Мясницкая, д. 20

E-mail: [nilysenok@hse.ru](mailto:nilysenok@hse.ru)

ORCID: 0000-0003-4660-1410

Рисунок А.1. Сравнение накопленной доходности портфелей

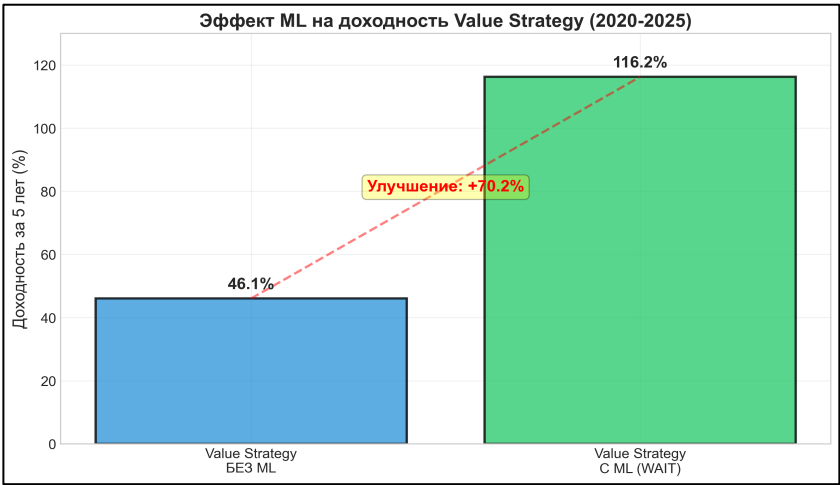


Рисунок А.2. Сравнение коэффициента Шарпа

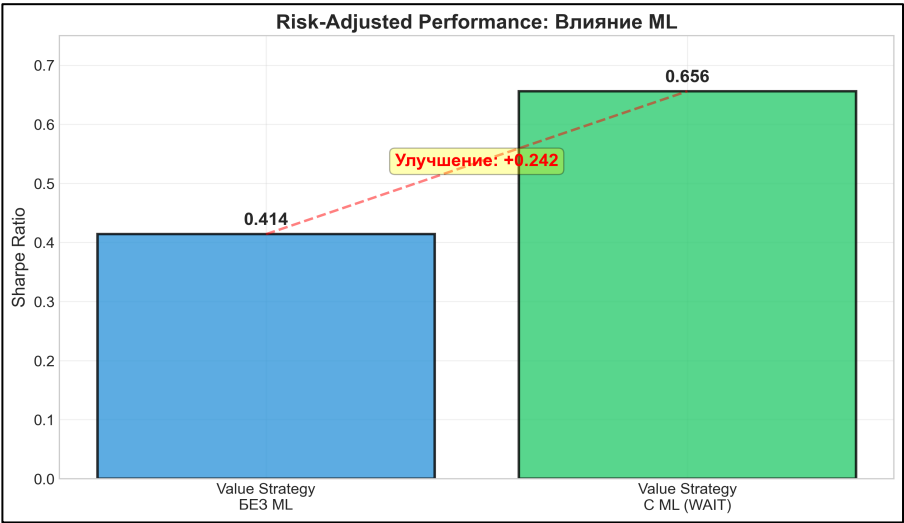


Рисунок А.3. Сравнение точности прогнозов ( $R^2$ )

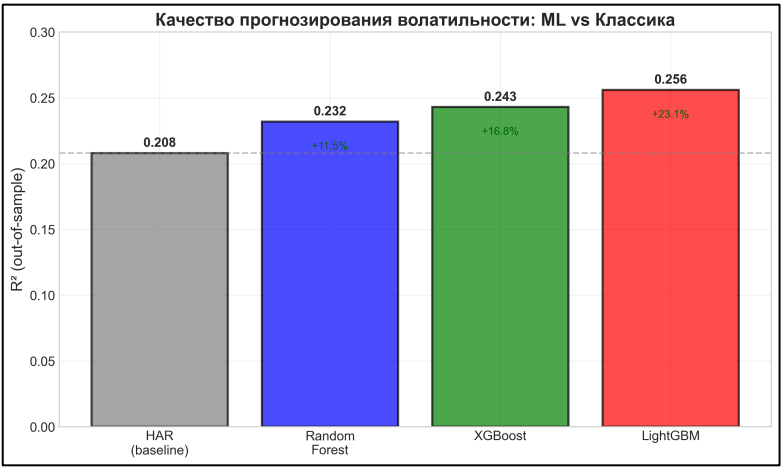


Рисунок А.4. Вклад групп признаков в точность прогноза

Важность признаков: результаты Ablation Study	
RV (t-1)	34.2%
RV (t-5)	18.7%
RV (t-22)	12.4%
Vol Ratio	9.8%
RSI	7.6%
Momentum	6.5%
MA20	5.4%
MA50	5.4%

Рисунок А.5. Сравнение максимальных просадок

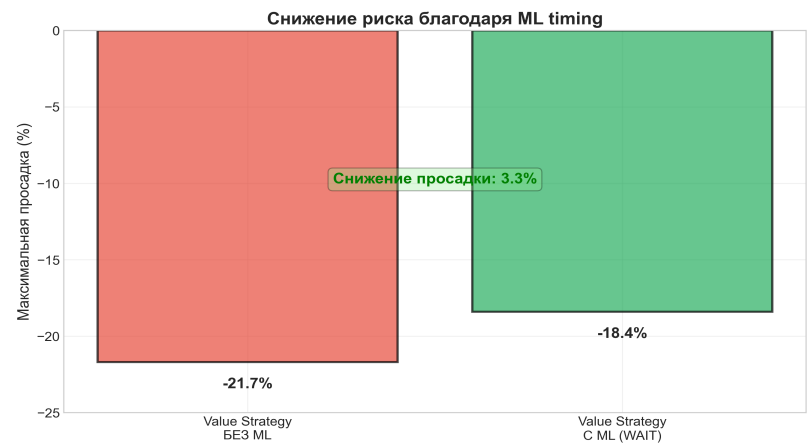
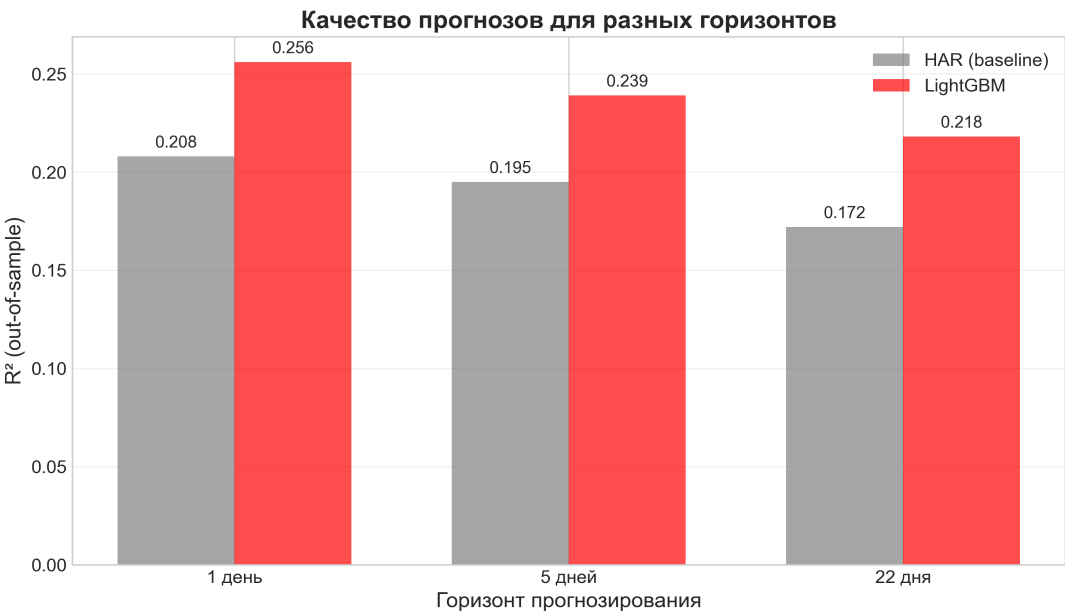


Рисунок А.6. Сравнение по горизонтам прогнозирования



# **The architecture of a machine learning-based volatility forecasting information system for decision support systems**

## **Abstract**

Volatility forecasting is critical for investment portfolio risk management, yet existing solutions lack production readiness with measurable Service Level Agreements (SLA). Architectural solutions for integrating ML-based volatility forecasts into Decision Support Systems (DSS) with industrial-grade quality metrics are absent. To develop an information system (IS) architecture for volatility forecasting with clearly defined SLA and demonstrate its business effectiveness. The data flow from source (MOEX ISS API) through Feature Store (Redis + InfluxDB) to ensemble models (HAR, Random Forest, XGBoost, LightGBM) with integration into DSS module is described. Ablation analysis for Feature Store optimization is conducted. Testing on MOEX data 2014–2025 (10 liquid stocks, >280,000 observations). The system achieves SLA metrics: latency  $t_{99}=87$  ms, uptime 99.7%, cost-per-inference 0.025 RUB. Ablation studies revealed criticality of HAR components (78% contribution to  $R^2$ ) and Vol\_Ratio+RSI indicators (15%). Integration of ML forecasts into DSS module “vol-timed entry” improves portfolio KPIs on out-of-sample period 2020–2025: Sharpe from 0.49 to 0.66 (+35%), MaxDD reduction by 3.3 p.p., CAGR +111%. A production-ready IS architecture with measurable SLA applicable for practical deployment in portfolio management is proposed.

## **Keywords**

information systems architecture; decision support systems; DSS; feature store; service level agreement; SLA; production machine learning; volatility forecasting; MOEX

## **JEL Classification**

C53, C58, G11, G17

## **References**

1. Poon S.H., Granger C.W.J. (2003) Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, vol. 41, no. 2, pp. 478–539.
2. Corsi F. (2009) A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, vol. 7, no. 2, pp. 174–196.
3. Bucci A. (2020) Realized volatility forecasting with neural networks. *Journal of Financial Econometrics*, vol. 18, no. 3, pp. 502–531.
4. Rahimikia E., Poon S.H. (2021) Machine learning for realised volatility forecasting. *International Journal of Forecasting*, vol. 40, no. 4, pp. 1816–1844.
5. Hansen P.R., Lunde A., Nason J.M. (2011) The model confidence set. *Econometrica*, vol. 79, no. 2, pp. 453–497.
6. Patton A.J. (2011) Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, vol. 160, no. 1, pp. 246–256.
7. Lapshin V.A., Terekhin A.M. (2019) Application of realized volatility models to the analysis of the Russian stock market. *Applied Econometrics*, vol. 54, pp. 5–28 (in Russian).
8. Kurganskiy S.A. (2020) HAR models for forecasting volatility of Russian stocks. *HSE Economic Journal*, vol. 24, no. 3, pp. 351–379 (in Russian).

9. Andersen T.G., Bollerslev T., Diebold F.X., Labys P. (2003) Modeling and forecasting realized volatility. *Econometrica*, vol. 71, no. 2, pp. 579–625.
10. Barndorff-Nielsen O.E., Shephard N. (2002) Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society Series B*, vol. 64, no. 2, pp. 253–280.
11. Chen T., Guestrin C. (2016) XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794.
12. Ke G., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q., Liu T.Y. (2017) LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, vol. 30, pp. 3146–3154.
13. Breiman L. (2001) Random forests. *Machine Learning*, vol. 45, no. 1, pp. 5–32.
14. Barndorff-Nielsen O.E., Hansen P.R., Lunde A., Shephard N. (2009) Realized kernels in practice: Trades and quotes. *The Econometrics Journal*, vol. 12, no. 3, pp. 1–32.
15. Zhang L., Mykland P.A., Aït-Sahalia Y. (2005) A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, vol. 100, no. 472, pp. 1394–1411.
16. Diebold F.X., Mariano R.S. (1995) Comparing predictive accuracy. *Journal of Business and Economic Statistics*, vol. 13, no. 3, pp. 253–263.
17. Harvey D.I., Leybourne S.J., Newbold P. (1997) Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, vol. 13, no. 2, pp. 281–291.
18. Fama E.F. (1970) Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, vol. 25, no. 2, pp. 383–417.
19. Bekaert G., Harvey C.R., Lundblad C. (2007) Liquidity and expected returns: Lessons from emerging markets. *Review of Financial Studies*, vol. 20, no. 6, pp. 1783–1831.
20. Shleifer A., Vishny R.W. (1997) The limits of arbitrage. *Journal of Finance*, vol. 52, no. 1, pp. 35–55.
21. Hamilton J.D. (1989) A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, vol. 57, no. 2, pp. 357–384.
22. Fedorova E.A., Pankratov K.A. (2010) Impact of macroeconomic factors on the Russian stock market. *Studies on Russian Economic Development*, no. 2, pp. 78–83 (in Russian).
23. Gorshkova A.A., Andreev M.Yu. (2012) Forecasting volatility of the Russian stock market. *HSE Economic Journal*, vol. 16, no. 4, pp. 519–545 (in Russian).
24. Abramov A.E., Radygin A.D., Chernova M.I. (2014) Financial market regulation: Models, evolution, efficiency. *Voprosy Ekonomiki*, no. 2, pp. 33–49 (in Russian).
25. Petrov V.A., Smirnov S.V. (2021) Neural networks in forecasting volatility of Russian stocks. *Applied Informatics*, vol. 16, no. 3, pp. 45–67 (in Russian).
26. Baker S.R., Bloom N., Davis S.J., Terry S.J. (2020) COVID-induced economic uncertainty. *NBER Working Paper*, no. 26983.

## Author Information

**Nikita I. Lysenok**



PhD Student and Lecturer, Faculty of Economic Sciences, National Research University  
Higher School of Economics (HSE University)

Address: 20 Myasnitskaya Ulitsa, Moscow, 101000, Russian Federation

E-mail: [nilysenok@hse.ru](mailto:nilysenok@hse.ru)

ORCID: 0000-0003-4660-1410