

УДК 338.27: 004.85 JEL C55

UDC 338.27: 004.85 JEL C55

5.2.2. Математические, статистические и инструментальные методы в экономике (физико-математические науки, экономические науки)

5.2.2. Mathematical, statistical and instrumental methods of economics (physical and mathematical sciences, economic sciences)

МАТЕМАТИЧЕСКИЕ МЕТОДЫ И ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА ПРОГНОЗИРОВАНИЯ СЛОЖНЫХ ПРОЦЕССОВ

MATHEMATICAL METHODS AND INSTRUMENTAL TOOLS FOR PREDICTING COMPLEX PROCESSES

Кумратова Альфира Менлигуловна
канд. экон. наук, Доцент
ФГБОУ ВО «Кубанский государственный аграрный университет имени И. Т. Трубилина»

Kumratova Alfira Menigulovna
Candidate of Economic Sciences, Associate Professor
Kuban State Agrarian University named after I. T. Trubilin

Розентул Илья Васильевич
Магистрант
Негосударственное аккредитованное некоммерческое частное образовательное учреждение высшего образования "Академия маркетинга и социально-информационных технологий-ИМСИТ"

Rosentul Ilya Vasilyevich
Undergraduate
Non-governmental accredited non-profit private educational institution of higher education "Academy of Marketing and Socio-Information Technologies-IMSIT"

Василенко Андрей Игоревич
Аспирант
ФГБОУ ВО «Кубанский государственный аграрный университет имени И. Т. Трубилина»

Vasilenko Andrey Igorevich
Graduate student
Kuban State Agrarian University named after I. T. Trubilin

В статье рассматривается использование различных математических и инструментальных средств для прогнозирования сложных процессов (на примере больших данных отдельных элементов финансового рынка). В работе представлена демонстрация авторского инструментального средства параллельного анализа Big Data на базе алгоритмов машинного обучения. Оценка эффективности полученных прогнозов значений отдельных элементов финансового рынка получена посредством реализации различных метрик ошибок. Актуальность темы исследования обусловлена растущей потребностью в высокоточных прогнозах для эффективного управления финансовыми рисками и принятия стратегических решений. Отличительной особенностью разработанной информационной системы является возможность прогнозирования временных рядов на заданное количество шагов вперед, а также построение графиков сравнения прогнозов для различных методов прогнозирования относительно фактических данных (язык программирования Java с использованием дополнительных библиотек, фреймворков и инструментальных средств, таких как Deeplearning4j, Apache Commons Math3, Apache POI, JavaFX, SceneBuilder, SLF4J, OpenCSV, Maven). Подбор гиперпараметров в системе реализован при помощи улучшенного

The article discusses the use of various mathematical and instrumental tools for predicting complex processes (using the example of big data of individual elements of the financial market). The paper presents a demonstration of the author's tool for parallel analysis of Big Data based on machine learning algorithms. The evaluation of the effectiveness of the obtained forecasts of the values of individual elements of the financial market was obtained through the implementation of various error metrics. The relevance of the research topic is due to the growing need for high-precision forecasts for effective financial risk management and strategic decision-making. A distinctive feature of the developed information system is the ability to predict time series for a given number of steps forward, as well as plotting comparison forecasts for various forecasting methods relative to actual data (Java programming language using additional libraries, frameworks and tools such as Deeplearning4j, Apache Commons Math3, Apache POI, JavaFX, SceneBuilder, SLF4J, OpenCSV, Maven). The selection of hyperparameters in the system is implemented using an improved Random Search algorithm. The essence of the algorithm is to generate random configurations of hyperparameters, it uses multithreading to train several models with different hyperparameters in parallel and selects the model with the best performance

алгоритма Random Search (случайного поиска). Суть алгоритма заключается в генерации случайных конфигурации гиперпараметров, он использует многопоточность для параллельного обучения нескольких моделей с различными гиперпараметрами и выбирает модель с лучшей производительностью

Ключевые слова: СТАТИСТИЧЕСКИЙ АНАЛИЗ, НЕЙРОННЫЕ СЕТИ, ПРОГНОЗИРОВАНИЕ, ВРЕМЕННОЙ РЯД, СЛОЖНЫЙ ПРОЦЕСС, БОЛЬШИЕ ДАННЫЕ

Keywords: STATISTICAL ANALYSIS, NEURAL NETWORKS, FORECASTING, TIME SERIES, COMPLEX PROCESS, BIG DATA

<http://dx.doi.org/10.21515/1990-4665-204-017>

Введение

В условиях растущей неопределенности и высокой динамики окружающей среды, разработка и применение эффективных математических и инструментальных средств прогнозирования становится жизненно необходимым для принятия обоснованных решений [1, 2].

Процесс прогнозирования основан на использовании исторических данных для создания моделей, способных предсказать будущие значения интересующих переменных, временного ряда. Эти модели могут быть как простыми линейными, так и сложными нелинейными, чтобы соответствовать особенностям анализируемых данных и требованиям к точности прогноза. Среди широкого спектра методов прогнозирования, которые также можно классифицировать на различные группы, наиболее часто применяются следующие:

1. Методы экспоненциального сглаживания:

– Простое экспоненциальное сглаживание: метод, который сглаживает временные ряды путем присвоения экспоненциально убывающих весов предыдущим наблюдениям. Он подходит для данных без ярко выраженной сезонности или трендов;

– Метод Хольта: расширение простого экспоненциального сглаживания, которое учитывает линейные тренды в данных, что

позволяет более точно прогнозировать временные ряды с трендовыми компонентами;

– Метод Хольта-Уинтерса: Он известен как метод экспоненциального сглаживания, представляет собой мощный инструмент для прогнозирования временных рядов, основывается на анализе сезонных колебаний и выявлении трендов, что позволяет эффективно предсказывать будущие значения. Суть метода заключается в использовании три параметра: уровня, тренда и сезонности, что дает возможность создать комплексную модель, учитывающую как долгосрочные, так и краткосрочные изменения. Каждый из параметров обновляется по мере поступления новых данных, что обеспечивает высокой степень адаптивности и точности прогнозов.

2. Авторегрессионные модели прогнозирования. Они представляют собой мощный инструмент для анализа временных рядов, позволяющий выявлять зависимости между текущими значениями и их предыдущими наблюдениями. Эти модели основываются на предположении, что текущее значение ряда можно выразить через линейную комбинацию его прошлых значений, что делает их особенно полезными в экономике, финансах и других науках. Существует несколько типов авторегрессионных моделей, включая $AR(p)$, где p обозначает количество лагов, используемых для предсказания. Основной процесс построения модели включает в себя идентификацию порядка модели с помощью критериев информационного выбора, а также оценку параметров. Значение авторегрессионного подхода заключается не только в возможности прогнозирования, но и в анализе динамики. Один из основных недостатков авторегрессионных моделей – это их чувствительность к неправильному выбору порядка и предположениям о стационарности данных. Перед применением таких моделей необходима тщательная предобработка данных и проверка на

стационарность, что позволяет обеспечить надежность сделанных прогнозов. В работе использованы алгоритмы моделей ARIMA и GARCH.

3. Модели на основе нейронных сетей представляют собой мощный инструмент в арсенале современных технологий. Современные достижения в области глубокого обучения, включающие сверточные и рекуррентные нейронные сети, открывают новые горизонты для различных приложений, от медицины до финансов. В частности, они превращают алгоритмы в «умные» системы, способные адаптироваться к изменениям и обучаться на своих ошибках, что в конечном итоге улучшает их производительность и эффективность. В работе апробированы алгоритмы моделей MLP (многослойные перцептроны) и LSTM (долгая краткосрочная память). Последние особенно полезны в задачах, где данные обладают сложной внутренней структурой и долгосрочной зависимостью.

Далее рассмотрим преимущества и недостатки методов прогнозирования.

Авторегрессионные модели (АР) представляют собой мощный инструмент для анализа временных рядов, позволяя эффективно прогнозировать значения на основе предыдущих наблюдений. Одним из основных преимуществ АР-моделей является их простота и удобство в интерпретации. Они позволяют выявлять временные зависимости и циклы в данных, что облегчает принятие решений на основе статистической информации. Апробация авторегрессионных моделей связано и с определёнными недостатками: 1) эти модели предполагают линейность зависимости, что не всегда отвечает реальной структуре данных; 2) АР-модели могут быть чувствительны к выбросам и несоответствиям, что может исказить результаты. При исследовании необходимо учитывать, что для успешного применения целесообразно предварительное тестирование на стационарность темпорального временного ряда. Отметим, что АР-

модели не учитывают экзогенные переменные, что может ограничивать их прогностическую силу в сложных системах. Выбор данной модели требует тщательного анализа и понимания специфики исследуемых данных.

Методы экспоненциального сглаживания являются популярным инструментом в предсказании временных рядов благодаря своей простоте и эффективности. Одним из основных преимуществ этих методов является их способность быстро адаптироваться к изменениям в данных, что делает их особенно полезными в условиях волатильности. Они также требуют минимального объема ретроспективных данных для корректного функционирования, что упрощает их применение в разных сценариях. Рассмотрим недостатки: 1) экспоненциальное сглаживание основывается на предположении о постоянности трендов и сезонности, что может привести к некорректным прогнозам в случаях резких изменений в паттернах данных (наличие/частота риск-экстремальных значений); 2) выбор параметров сглаживания требует тщательной настройки, что может создать дополнительные сложности для аналитиков. При проведении исследования нужно учитывать ограничения методов экспоненциального сглаживания и применять их с осторожностью, особенно в условиях неопределенности и непредсказуемости данных.

Модели на основе нейронных сетей получили широкое признание в последние годы благодаря своим впечатляющим результатам в различных областях, таких как обработка изображений, естественный язык и даже игры. Отличительной особенностью нейронных сетей является их способность самостоятельно выявлять сложные паттерны в данных, что позволяет достигать высокой точности при распознавании и классификации. Они также очень масштабируемы и могут обрабатывать огромные объемы данных, что делает их подходящими для задач, требующих учета множества факторов. Далее перечислим недостатки таких моделей: 1) необходимость в значительном количестве обучающих

данных, что может стать проблемой в случаях, когда данные трудно собрать или они имеют низкое качество; 2) нейронные сети часто рассматриваются как «черные ящики», что затрудняет интерпретацию результатов и понимание внутренней механики принятия решений; 3) высокие требования к характеристикам технических средств для обработки больших объемов данных. При выборе нейронной сети необходимо балансировать между их преимуществами и недостатками, учитывая специфику задач [3, 4, 5].

В рамках представленного исследования будет продемонстрирована авторская информационная система параллельного анализа больших данных «Finam Prophet». Это настольное (desktop) приложение, разработанная на языке программирования Java с использованием дополнительных библиотек, фреймворков и инструментальных средств, таких как Deeplearning4j, Apache Commons Math3, Apache POI, JavaFX, SceneBuilder, SLF4J, OpenCSV, Maven.

В качестве объектов исследования будут выступать временные ряды, сформированные на основе данных финансового рынка, а именно:

- котировки акций компании ООО «Яндекс» за период 01.01.2020 – 16.01.2024 с периодичностью в 1 день.
- котировки акций компании ПАО «Лукойл» за период 01.01.2020 – 16.01.2024 с периодичностью в 1 день.

Графики финансовых временных рядов будут представлены на рисунках 1–2.

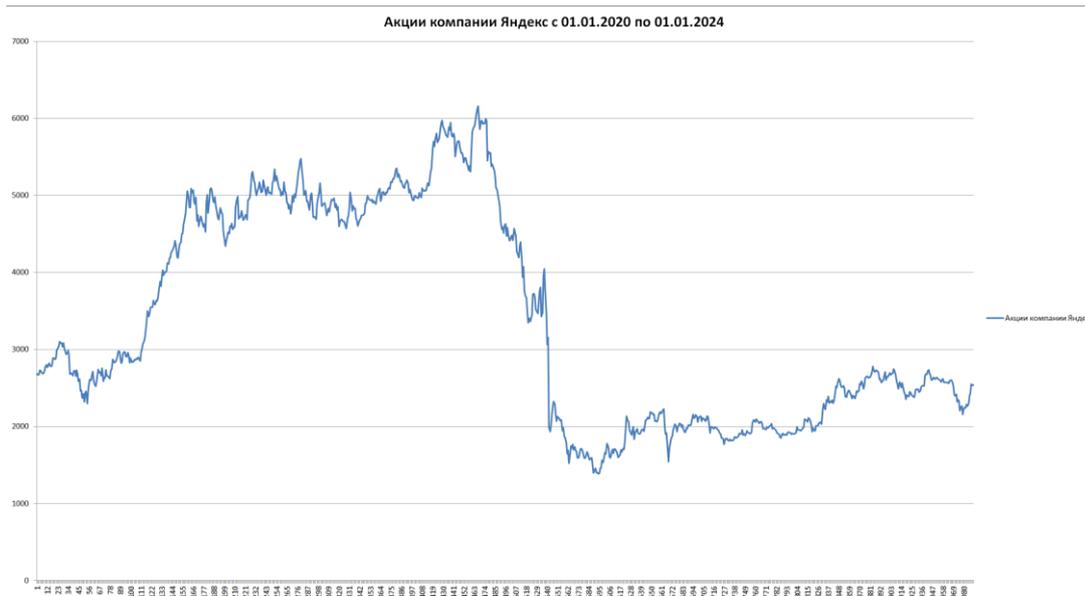


Рисунок 1 – График котировок значений цен на акции компании ООО «Яндекс»

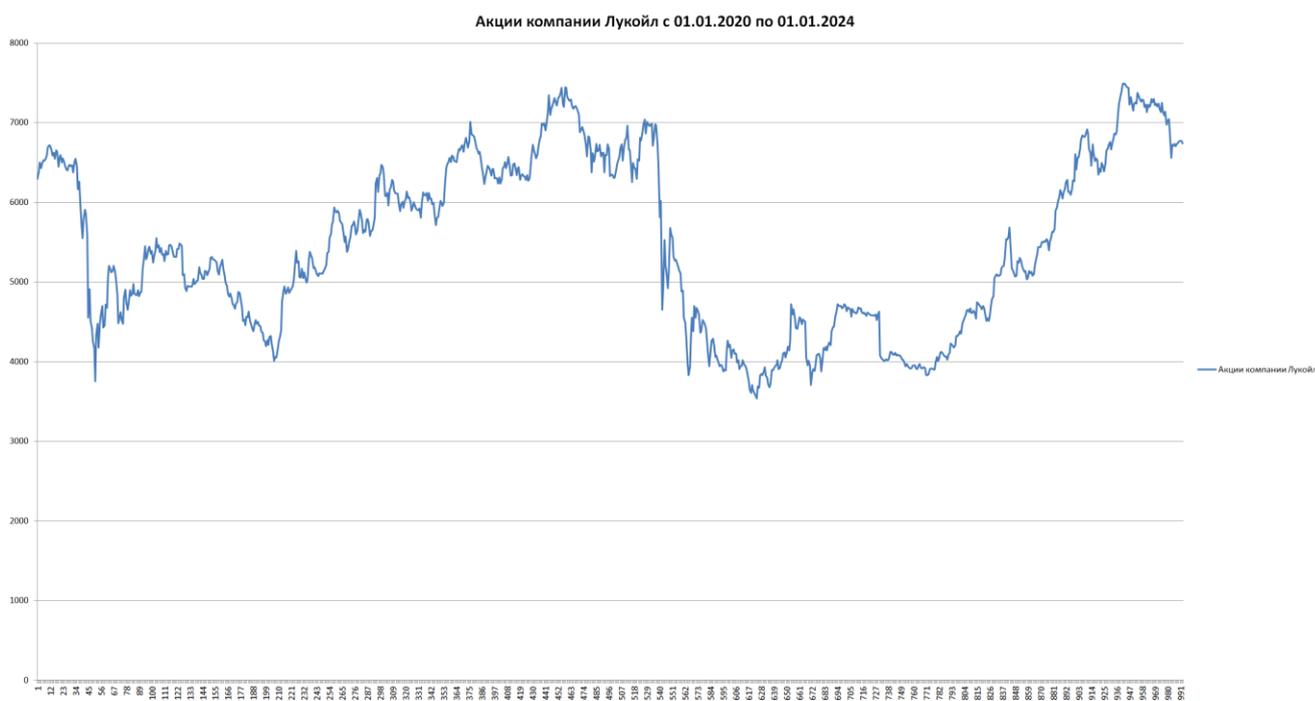


Рисунок 2 – График котировок значений цен на акции компании ПАО «Лукойл»

Также отобразим выборку фактических значений на следующие 10 шагов временных рядов, сформированную за период с 01.01.2024 по 16.01.2024 (рисунки 3 и 4).

Фактические значения от 01.01.2024 по 16.01.2024

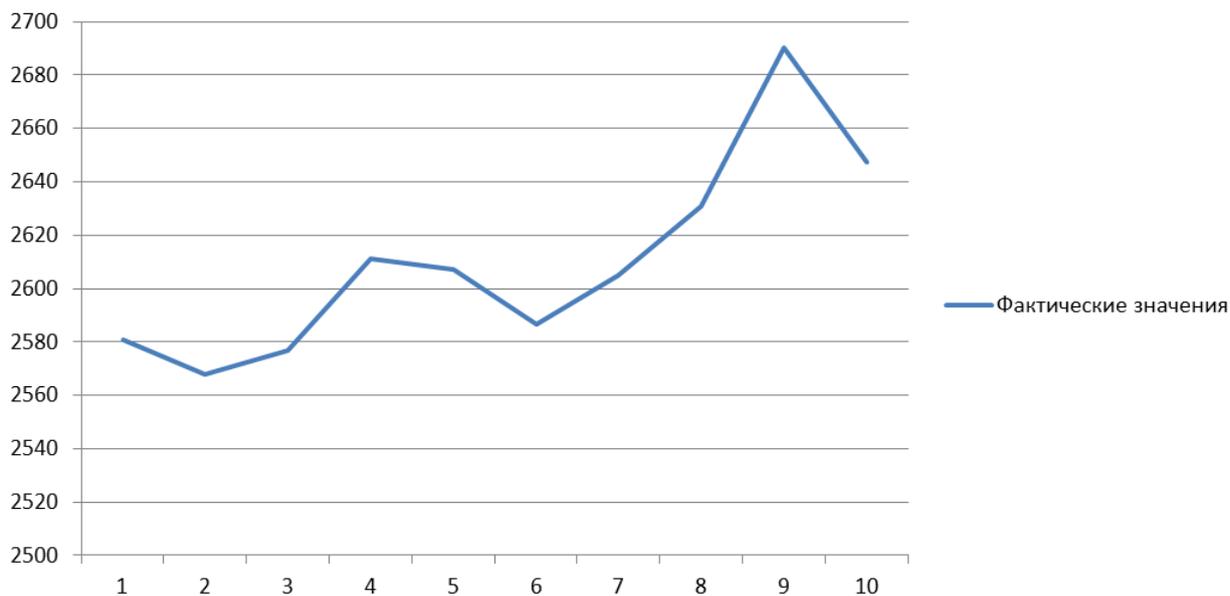


Рисунок 3 – Фактические значения котировок акций ООО «Яндекс» на следующие 10 шагов временного ряда

Фактические значения от 01.01.2024 по 16.01.2024

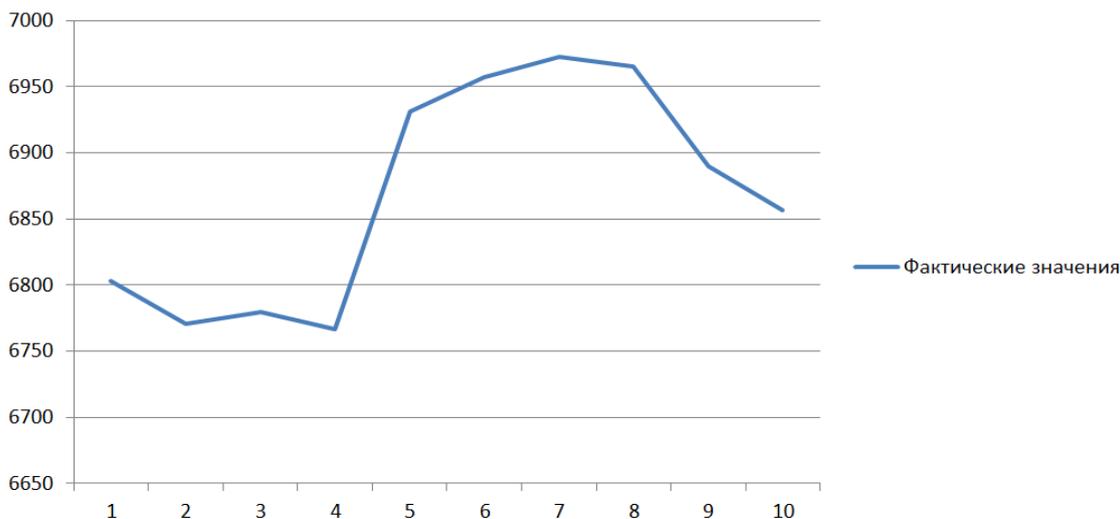


Рисунок 4 – Фактические значения котировок акций ПАО «Лукойл» на следующие 10 шагов временного ряда

Для построения прогнозов будем использовать экономико-математические методы, описанные выше. Для демонстрации работы программы, спрогнозированные системой значения котировок акций компании ООО «Яндекс», будут отображены на рисунках 5–11.

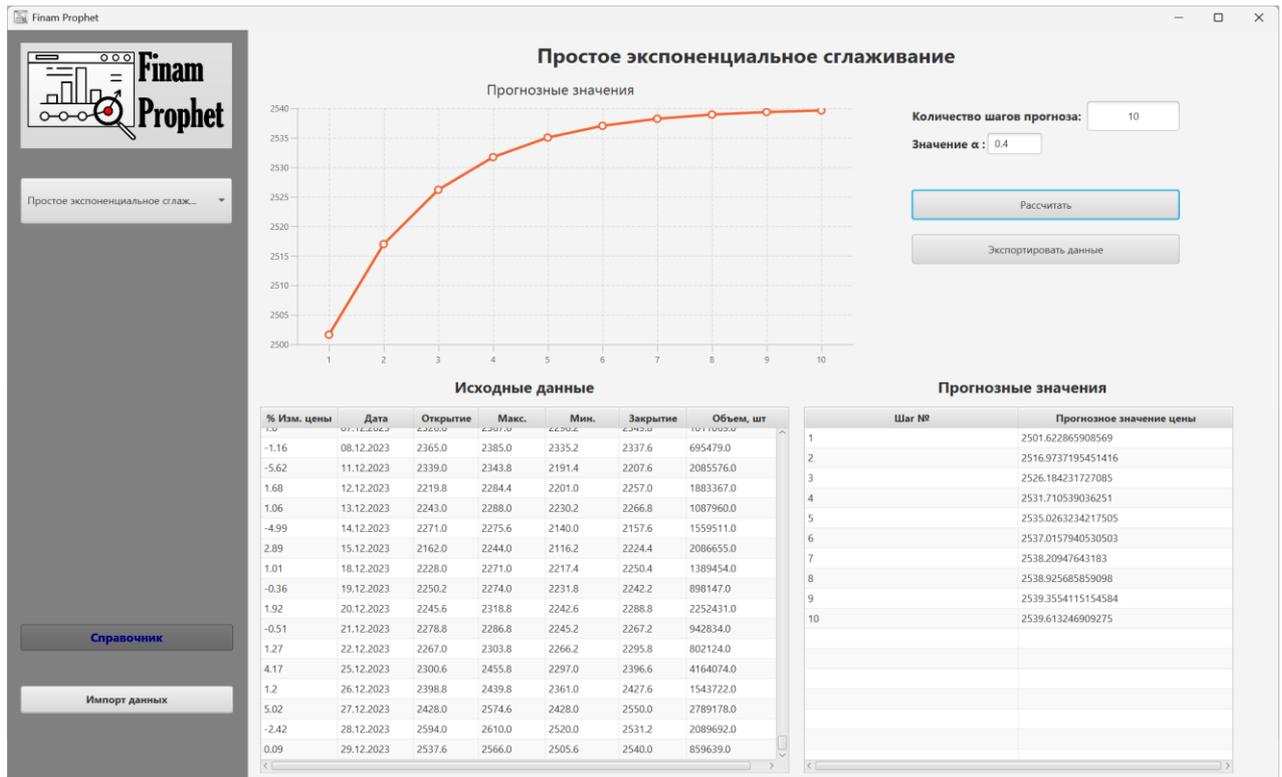


Рисунок 5 – Прогноз значений цен на акции ООО «Яндекс» методом простого экспоненциального сглаживания

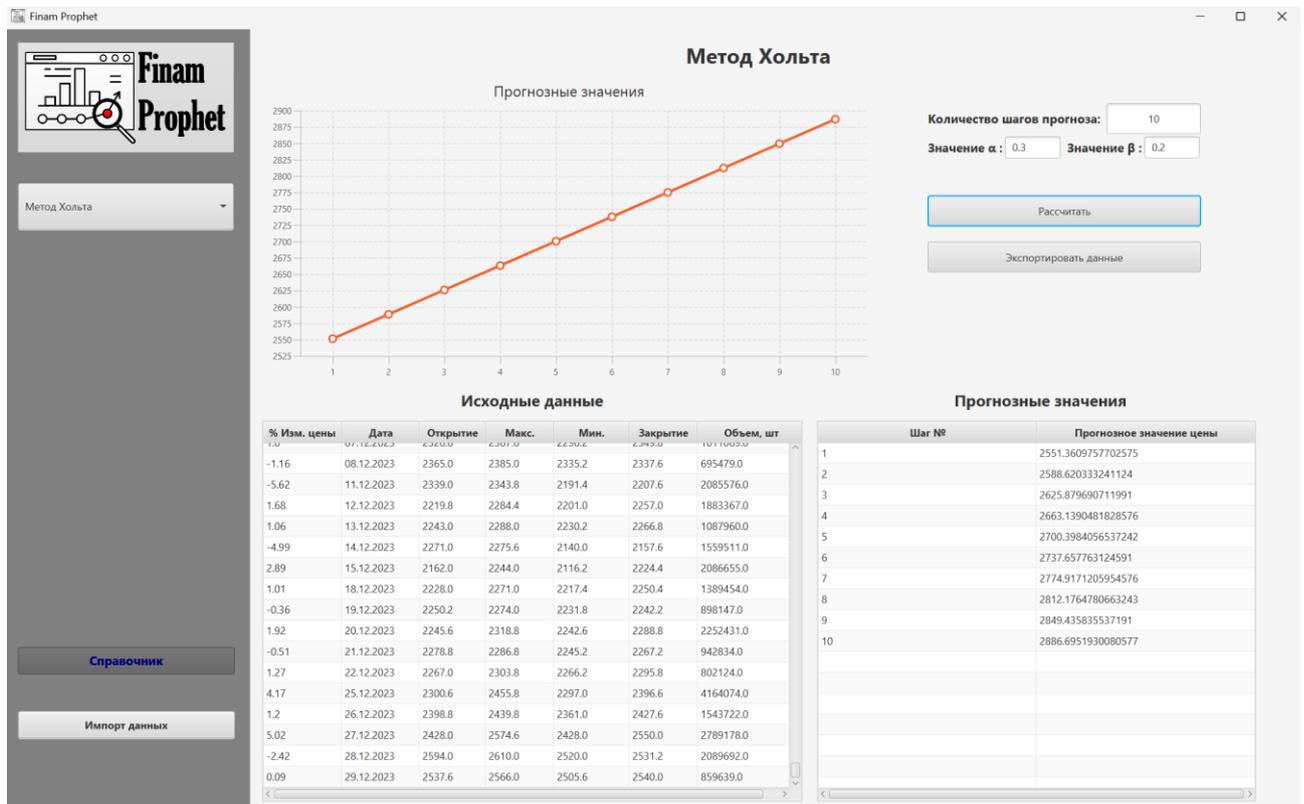


Рисунок 6 – Прогноз значений цен на акции ООО «Яндекс» методом Хольта

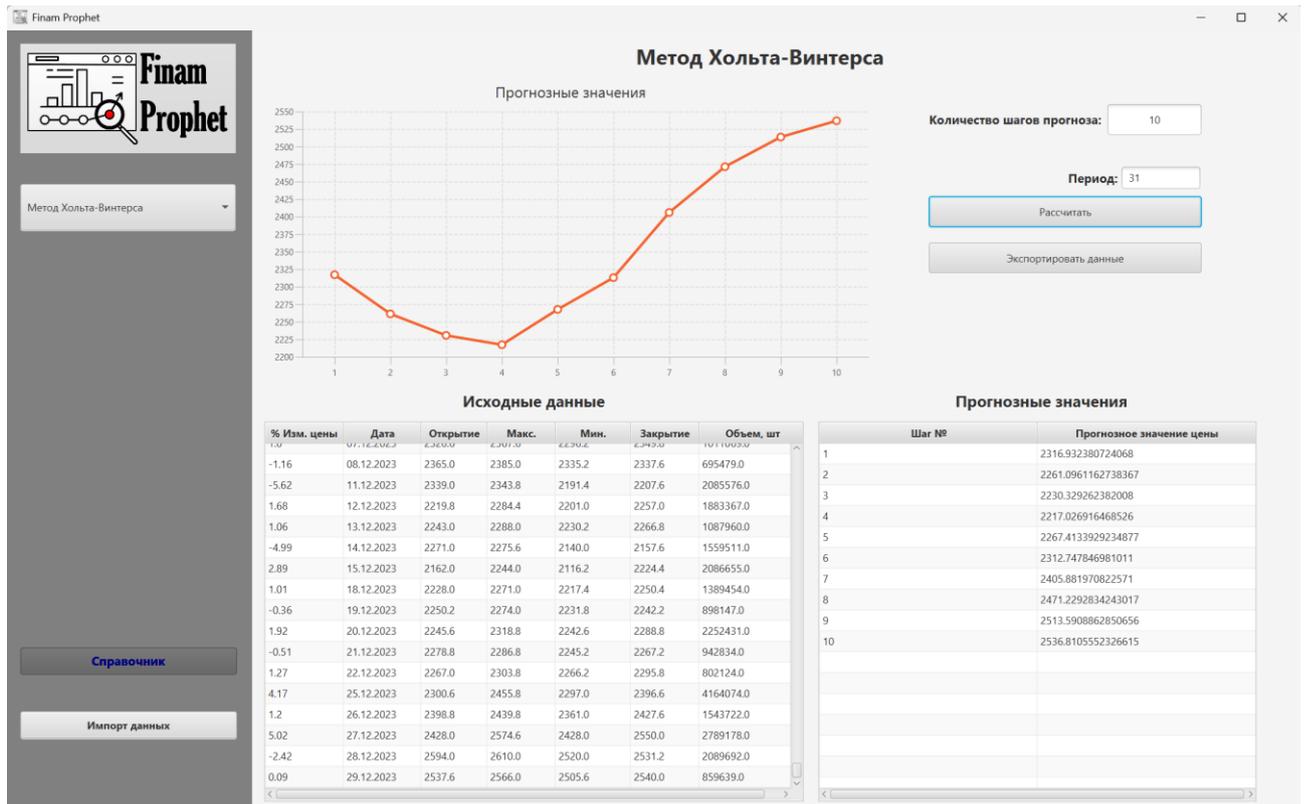


Рисунок 7 – Прогноз значений цен на акции ООО «Яндекс» методом Хольта-Уинтерса

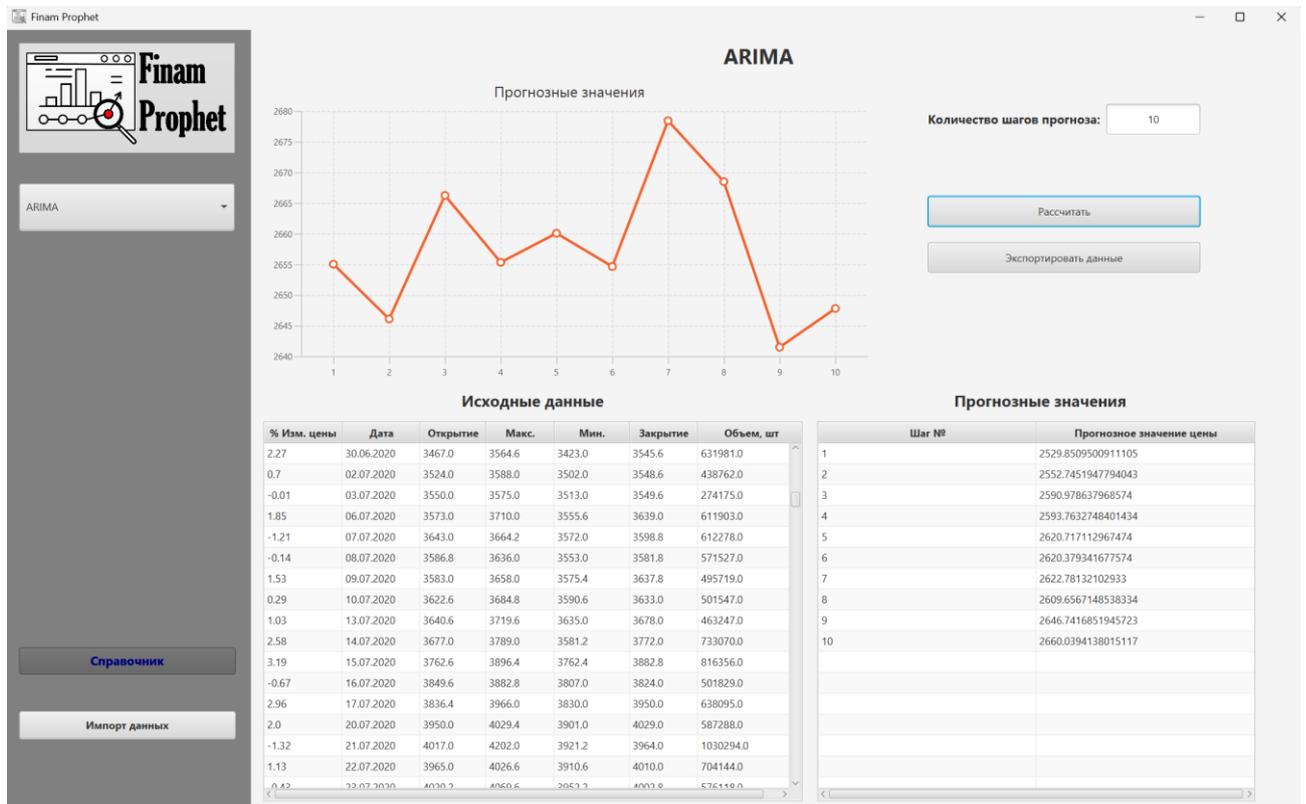


Рисунок 8 – Прогноз значений цен на акции ООО «Яндекс» методом ARIMA

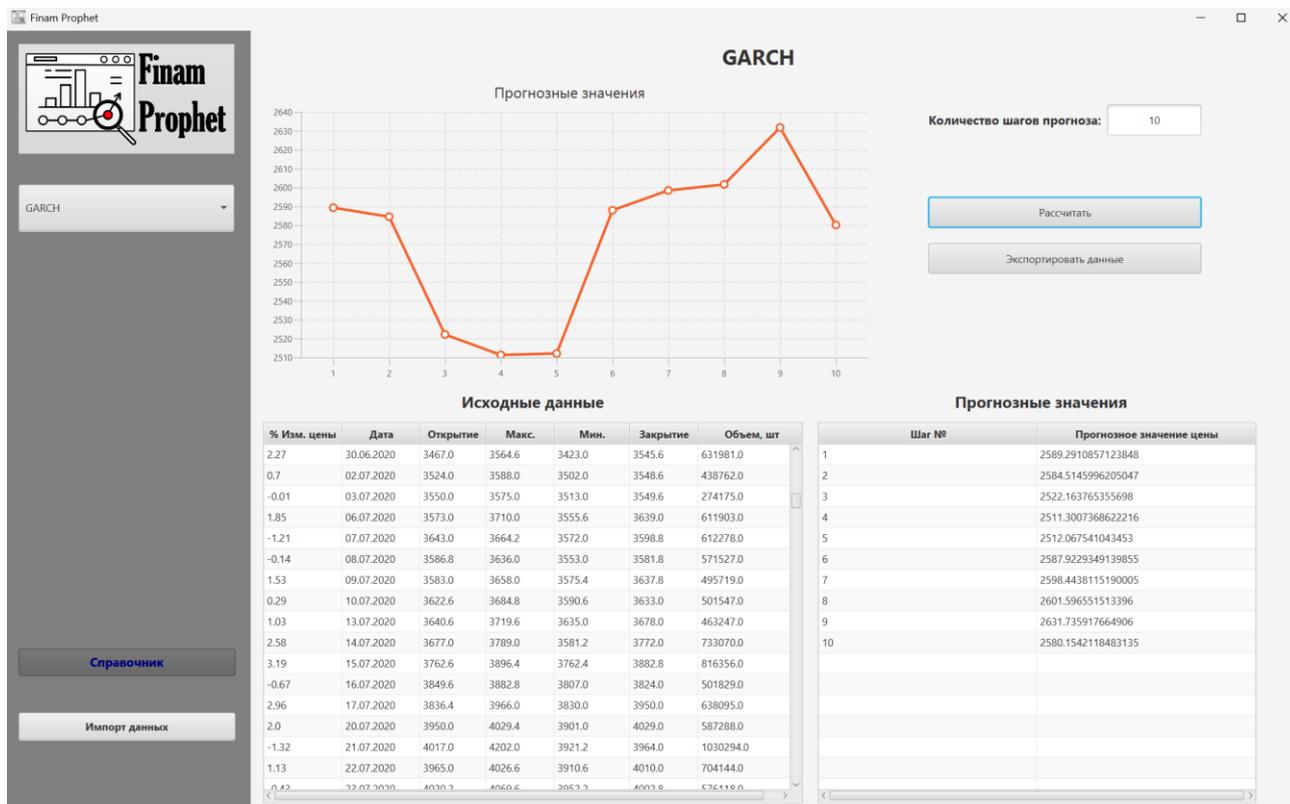


Рисунок 9 – Прогноз значений цен на акции ООО «Яндекс» методом GARCH

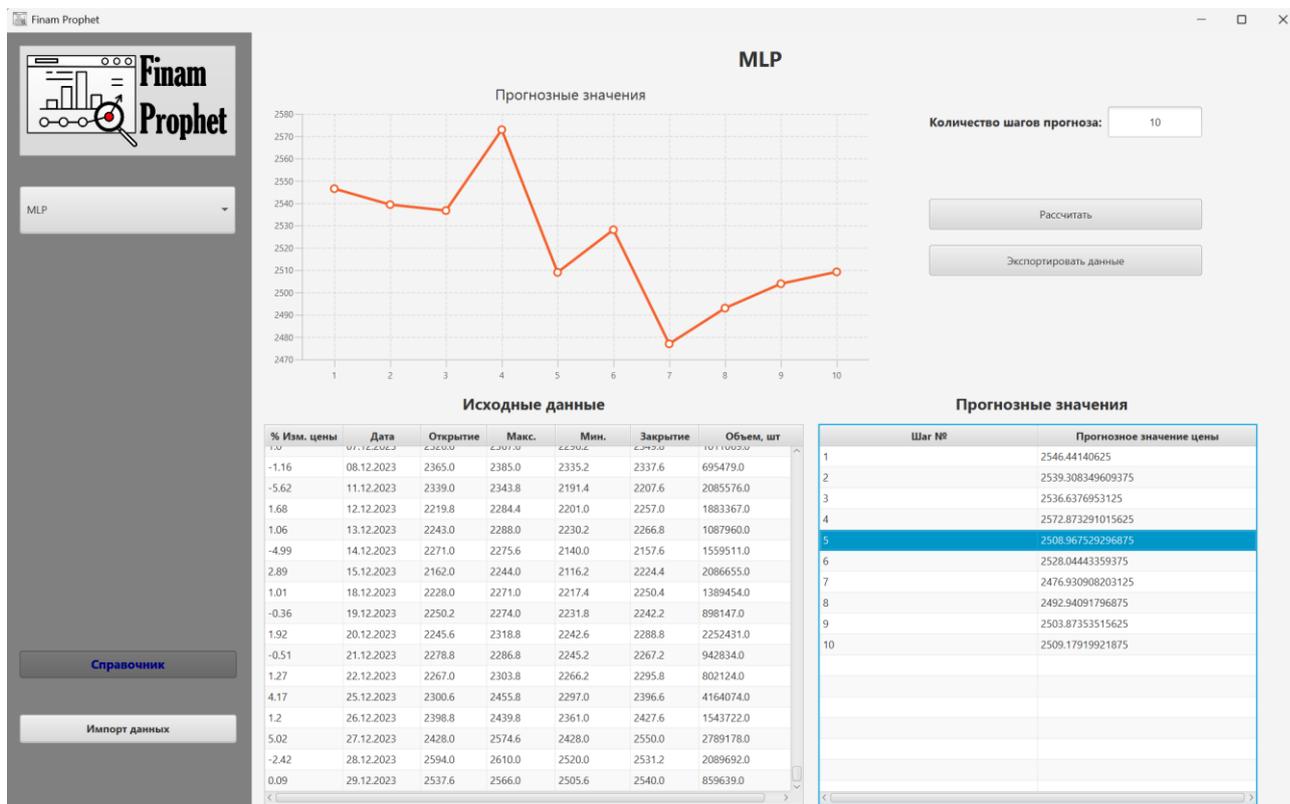


Рисунок 10 – Прогноз значений цен на акции ООО «Яндекс» методом MLP

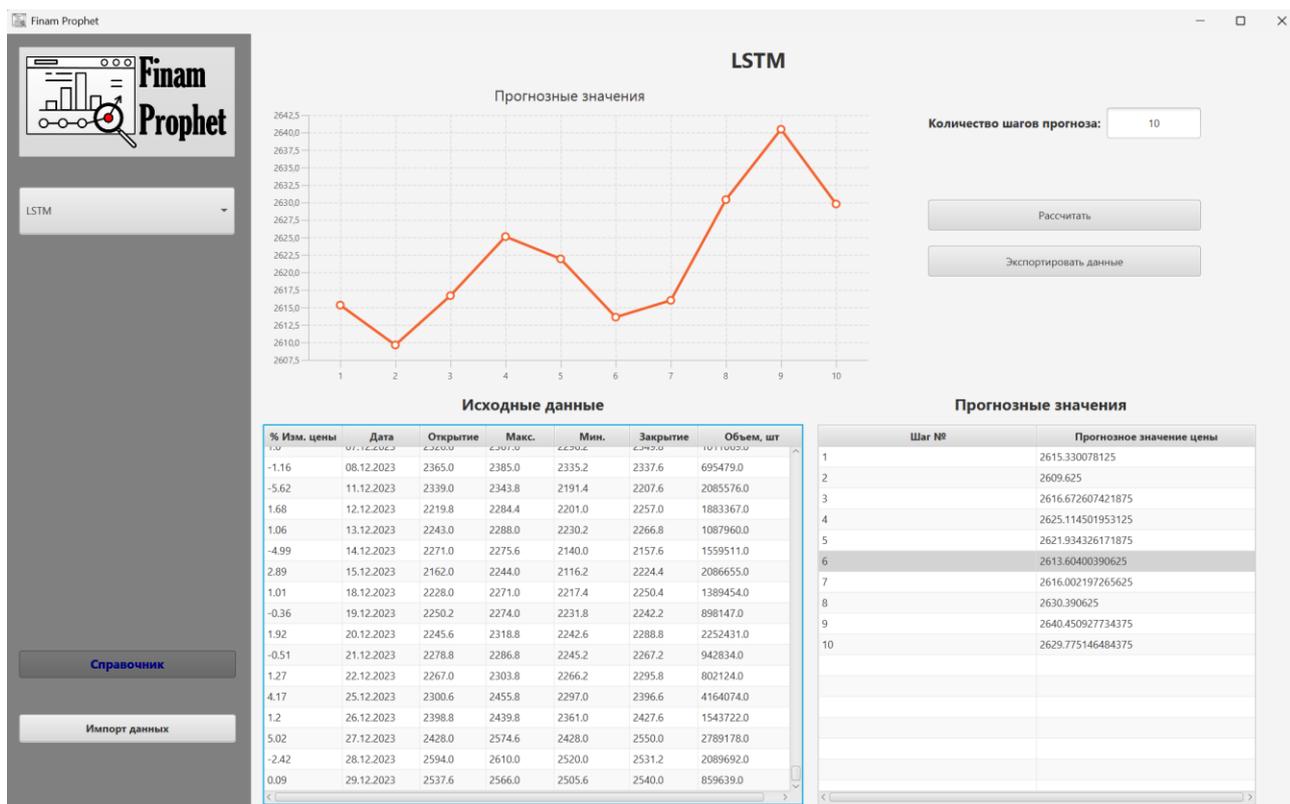


Рисунок 11 – Прогноз значений цен на акции ООО «Яндекс» методом LSTM

Аналогичные прогнозы будут сделаны и для акций компании ПАО «Лукойл». Далее представлены результаты сравнительного анализа прогнозных значений различных экономико-математических методов, с фактическими данными (рисунки 12 и 13).

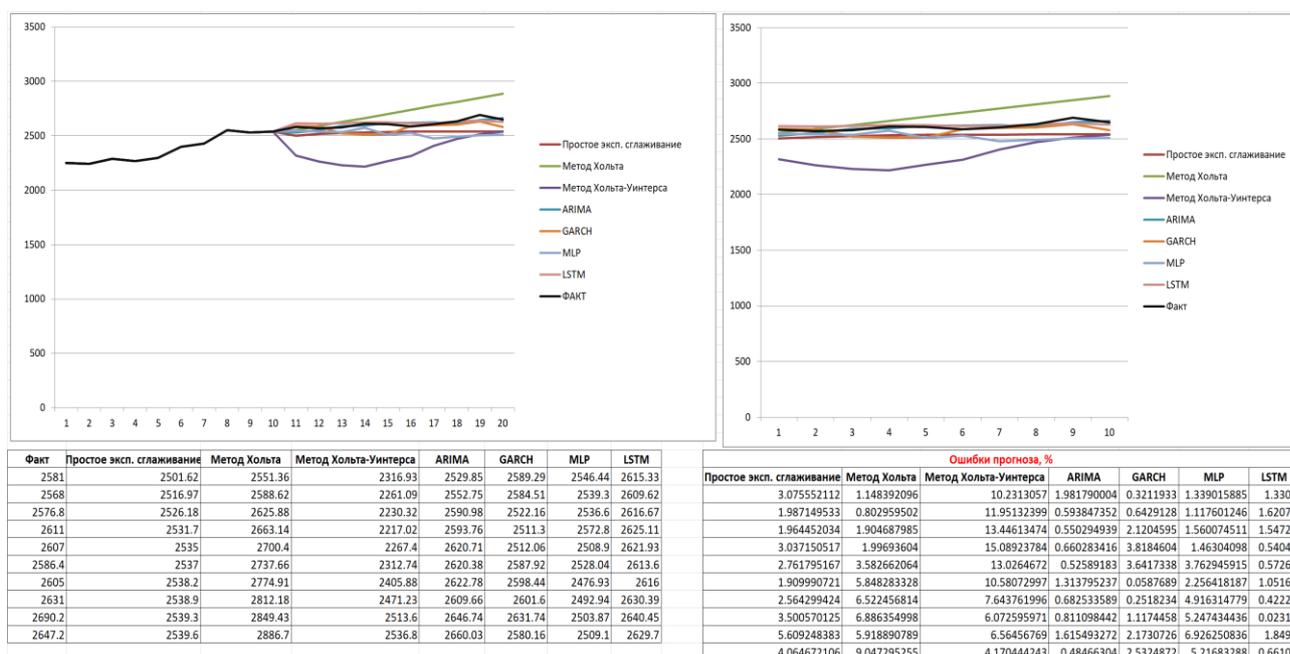


Рисунок 12 – Сравнение прогнозов котировок акций компании ООО «Яндекс»



Рисунок 13 – Сравнение прогнозов котировок акций компании ПАО «Лукойл»

Результаты сравнения прогнозов при помощи различных метрик ошибки будут представлены в таблицах 1 и 2.

Таблица 1 – Метрики ошибок для акций компании ООО «Яндекс»

Метрика	Простое эксп. сглаживание	Метод Хольта	Метод Хольта-Уинтерса	ARIMA	GARCH	MLP	LSTM
MAE	79.913	114.596	257.059	24.092	43.706	88.868	25.092
MSE	7267.37	18244.433	73796.425	751.468	3110.216	10737.583	856.491
RMSE	85.249	135.072	271.655	27.413	55.769	103.622	29.266
MAPE	3.047	4.366	9.878	0.922	1.668	3.381	0.962
MPE	3.047	-4.136	9.878	0.211	1.463	3.381	-0.455
R ²	-4.789	-13.532	-57.78	0.401	-1.477	-7.553	0.318

Таблица 2 – Метрики ошибок для акций компании ПАО «Лукойл»

Метрика	Простое эксп. сглаживание	Метод Хольта	Метод Хольта-Уинтерса	ARIMA	GARCH	MLP	LSTM
MAE	79.913	114.596	257.059	24.092	43.706	88.868	25.092
MSE	7267.37	18244.433	73796.425	751.468	3110.216	10737.583	856.491
RMSE	85.249	135.072	271.655	27.413	55.769	103.622	29.266
MAPE	3.047	4.366	9.878	0.922	1.668	3.381	0.962
MPE	3.047	-4.136	9.878	0.211	1.463	3.381	-0.455
R ²	-4.789	-13.532	-57.78	0.401	-1.477	-7.553	0.318

На основе полученных результатов можно сделать следующие выводы:

1. Отсутствие универсального метода прогнозирования: Даже несмотря на итоговые показатели, нет универсального метода прогнозирования, который был бы лучшим для всех наборов данных. Для акций компании ООО «Яндекс» лучшим методом оказался ARIMA, а для акций компании ПАО «Лукойл» – метод Хольта.

2. Зависимость от данных: Эффективность каждого метода зависит от конкретных данных, с которыми он работает. Те же методы могут показать совершенно разные результаты на других наборах данных. Например, метод Хольта-Уинтерса, который показал себя плохо в данном анализе, может оказаться более эффективным при прогнозировании других временных рядов.

3. Разнообразие методов: Различные методы имеют свои преимущества и недостатки. Модель LSTM продемонстрировала хорошие результаты в обеих выборках данных, что говорит о ее высокой универсальности. Средние результаты показала модель GARCH.

4. Необходимость адаптации методов: При выборе метода прогнозирования важно учитывать специфику данных и адаптировать подход к каждой конкретной ситуации. Даже если метод показал себя хорошо в одном случае, это не гарантирует его успех в другом.

В заключение, для достижения наилучших результатов в прогнозировании рекомендуется использовать несколько методов и проводить их сравнительный анализ на конкретных данных, чтобы выбрать наиболее подходящий для каждой ситуации.

Список использованных источников

1. Кумратова, А. М. Концептуальная основа получения и исследования максимального времени прогноза с заранее заданной точностью / А. М. Кумратова // Современная экономика: проблемы и решения. – 2017. – № 6(90). – С. 23-31.
2. Кумратова, А. М. Теория и практика моделирования, анализа и прогнозирования эволюционных социально-экономических систем методами нелинейной динамики / А. М. Кумратова // Труды Кубанского государственного аграрного университета. – 2017. – № 69. – С. 30-35.
3. Андрианова Е.Г., Чукалина Е.Р. Сравнение методов прогнозирования финансовых временных рядов // ИТ-Стандарт. 2021. № 2 (27). С. 40-45.. - 2021. - №2. - С. 40-45.
4. Дауб И. С. Исследование статистических методов прогнозирования временных рядов с трендом и сезонностью // StudNet. - 2021. - №5. - С. 1-23.
5. Временные ряды. Режим доступа: <https://rusforexclub.com/articles/24-finansovaya-matematika/99-vremennye-#c1> (дата обращения: 02.09.2024).

References

1. Kumratova, A. M. Konceptual'naya osnova polucheniya i issledovaniya maksimal'nogo vremeni prognoza s zaranee zadannoj tochnost`yu / A. M. Kumratova // Sovremennaya e`konomika: problemy` i resheniya. – 2017. – № 6(90). – S. 23-31.
2. Kumratova, A. M. Teoriya i praktika modelirovaniya, analiza i prognozirovaniya e`volyucionny`x social`no-e`konomicheskix sistem metodami nelinejnoj dinamiki / A. M. Kumratova // Trudy` Kubanskogo gosudarstvennogo agrarnogo universiteta. – 2017. – № 69. – S. 30-35.
3. Andrianova E.G., Chukalina E.R. Sravnenie metodov prognozirovaniya finansovy`x vremenny`x ryadov // IT-Standart. 2021. № 2 (27). S. 40-45.. - 2021. - №2. - S. 40-45.
4. Daub I. S. Issledovanie statisticheskix metodov prognozirovaniya vremenny`x ryadov s trendom i sezonnost`yu // StudNet. - 2021. - №5. - S. 1-23.
5. Vremenny`e ryady`. Rezhim dostupa: <https://rusforexclub.com/articles/24-finansovaya-matematika/99-vremennye-#c1> (data obrashheniya: 02.09.2024).