

## ОЦЕНКА ПРИМЕНИМОСТИ СТАТИСТИЧЕСКИХ И МАШИННЫХ МОДЕЛЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ИНФЛЯЦИИ

Астраханцева И.А., Герасимов А.С., Смирнова О.П.

Ирина Александровна Астраханцева\* (ORCID 0000-0003-2841-8639)

Герасимов Александр Сергеевич (ORCID 0000-0001-7867-2710)

Смирнова Ольга Павловна (ORCID 0000-0001-6563-9452)

Ивановский государственный химико-технологический университет,

Шереметевский проспект, д.7. Иваново, 153000, Россия

E-mail: smirnowa07@rambler.ru, i.astrakhantseva@mail.ru\*, aleksandr132542255@gmail.com

*Прогнозирование инфляции является критически важной задачей, которая играет значимую роль в формировании стратегий финансового планирования, инвестиционной деятельности и политики центральных банков. В условиях постоянно меняющейся экономической среды, способность точно предсказывать изменения уровня цен становится неотъемлемым инструментом для эффективного управления экономическими рисками.*

*В данной работе основное внимание уделяется разработке, анализу и сравнению различных подходов к прогнозированию инфляции, включая как современные методы машинного обучения, так и традиционные эконометрические модели.*

*В ходе исследования были разработаны и протестированы различные модели прогнозирования, включая градиентный бустинг, случайный лес, LSTM, сверточные нейронные сети (CNN), а также эконометрические методы, такие как ARIMA, VAR и Prophet. Проводилось обучение этих моделей для прогнозирования индекса потребительских цен (CPI) России на периоды в 1, 2 и 3 месяца вперед, а затем анализ и сравнение их эффективности. Это позволило выявить наиболее точные и эффективные подходы для решения задачи прогнозирования инфляции в России. Результаты данного исследования могут быть использованы специалистами в области экономики и финансов.*

**Ключевые слова:** прогноз инфляции, индекс потребительских цен, модели машинного обучения, эконометрические модели, градиентный бустинг, случайный лес, LSTM, CNN, Prophet, ARIMA и VAR.

## EVALUATING STATISTICAL AND MACHINE LEARNING MODELS FOR INFLATION FORECASTING

I.A. Astrakhantseva, A.S. Gerasimov, O.P. Smirnova

Irina A. Astrakhantseva\*(ORCID 0000-0003-2841-8639)

Alexandr S. Gerasimov (ORCID 0000-0001-7867-2710)

Olga P. Smirnova (ORCID 0000-0001-6563-9452)

Ivanovo state university of chemistry and technology, Sheremetev avenue, 7, Ivanovo, 153000, Russia

E-mail: i.astrakhantseva@mail.ru\*, aleksandr132542255@gmail.com, smirnowa07@rambler.ru

*Inflation forecasting is a critically important task that plays a significant role in shaping financial planning strategies, investment activities, and central bank policies. In an ever-changing economic environment, the ability to accurately predict price level changes is becoming an essential tool for effective economic risk management.*

*This paper focuses on the development, analysis and comparison of various approaches to inflation forecasting, including both modern machine learning methods and traditional econometric models. During the research, various forecasting models were developed and tested, including gradient boosting, random forest, LSTM, convolutional neural networks (CNN), as well as econometric methods such as ARIMA, VAR and Prophet. These models were trained to predict the consumer price index (CPI) of Russia for periods of 1, 2 and 3 months ahead, and then analyzed and compared their*

*effectiveness. This made it possible to identify the most accurate and effective approaches to solving the problem of forecasting inflation in Russia. The results of this study can be used by specialists in the field of economics and finance.*

**Keywords:** inflation forecast, consumer price index, machine learning models, econometric models, gradient boosting, random forest, LSTM, CNN, Prophet, ARIMA and VAR.

### ВВЕДЕНИЕ

Актуальность темы исследования обусловлена тем, что инфляция – это один из важнейших макроэкономических показателей, который оказывает значительное влияние на экономическую стабильность и развитие страны. Точное прогнозирование инфляции имеет огромное значение для принятия обоснованных решений в области финансов, инвестиций, политики центрального банка и многих других сферах. С развитием информационных технологий и методов машинного обучения появились новые возможности для улучшения точности прогнозирования инфляции. Внедрение ИИ (искусственного интеллекта) в процесс прогнозирования инфляции будет способствовать повышению прозрачности, объективности и обоснованности решений, принимаемых регулирующими органами. Их аргументация будет основываться не только на экспертных суждениях, но и на четких математических моделях, доступных для анализа и верификации [1], [2].

В то же время, с научной точки зрения, вопросы создания и применения искусственного интеллекта в области экономического прогнозирования относятся к числу недостаточно разработанных в виду ограничения данных, сложности моделирования нелинейных взаимосвязей, требования постоянного мониторинга и адаптации моделей к изменяющимся экономическим условиям, ограничения вычислительных ресурсов.

Инфляция представляет собой устойчивый рост общего уровня цен в экономике, и для ее количественной оценки и прогнозирования необходимы соответствующие единицы измерения. Индексы цен, такие как индекс потребительских цен (ИПЦ), индекс цен производителей и другие, служат именно этой цели [3]. В нашей работе был использован индекс ИПЦ.

Данный индекс отслеживает динамику цен на определенный набор товаров и услуг, представляющий типичную потребительскую корзину домохозяйства. В России с 1991 года принято рассчитывать этот показатель по методике Э.Ласпейреса. Именно ИПЦ является основным инструментом измерения инфляции [4]. Федеральная служба статистики вычисляет индексы цен для потребителей, принимая во внимание расходы домохозяйств на определенные товары.

В ее таблицах, представлены процентные расчеты инфляции на основе сравнения "месяц к месяцу" [5], [6].

Инфляция и индекс потребительских цен – тесно связанные, но не тождественные экономические понятия. Инфляция представляет собой устойчивый рост общего уровня цен в экономике, приводящий к снижению покупательной способности денег. Это макроэкономический процесс, затрагивающий всю экономику в целом.

ИПЦ рассчитывается на основе изменения цен фиксированной потребительской корзины товаров и услуг за определенный период (обычно месяц или год). Состав этой корзины отражает структуру потребления среднестатистического домохозяйства [7]. Темп инфляции приблизительно равен приросту ИПЦ в процентах минус 100%.

### МЕТОДЫ И МАТЕРИАЛЫ

В ходе данного исследования были разработаны такие модели прогнозирования, как градиентный бустинг, случайный лес, LSTM, сверточные нейронные сети (CNN), а также эконометрические методы, такие как ARIMA, VAR и Prophet. Проводилось обучение и тестирование этих моделей для прогнозирования ИПЦ на периоды в 1, 2 и 3 месяца вперед. Для обучения использовались данные (перечислены ниже) с января 2009 года по декабрь 2022 года, и последующая их проверка на тестовом наборе данных с января 2023 по март 2024 года. Далее был проведен анализ и сравнение эффективности различных методов прогнозирования, что позволило выявить наиболее точные и эффективные подходы для решения задачи прогнозирования инфляции в России.

Программный код для моделей был написан на языке Python. Для упрощения разработки использовались библиотеки Matplotlib, NumPy, Pandas, PyTorch, Prophet, Scikit-learn, Statsmodels. Для сравнения эффективности моделей была разработана базовая модель, которая предполагает, что значение инфляции в прогнозируемом месяце будет равно значению инфляции за последний известный месяц. Научная новизна заключается в уникальности результатов, полученных при комплексном анализе эффективности используемых методов прогнозирования инфляции в России, который для такого сочетания моделей был осуществлен впервые.

Практическая значимость исследования включает в себя создание новых инструментов и методов для прогнозирования инфляции. Интеграция таких технологий и методов обеспечивает возможность эффективного анализа больших объемов данных и выявления сложных зависимостей, что может повысить точность прогнозов и улучшить качество принимаемых экономических решений [8], [9].

Для сбора и построения моделей использовались открытые данные с официальных сайтов Росстата, Банка России, Московской биржи и других источников за период с января 2009 года по март 2024 года включительно.

Данные также включают в себя следующие экономические показатели: индекс потребительских цен (CPI) [39], среднемесячная фактическая ставка по кредитам (в руб.), предоставленным банками (MIACR) [10], цена на нефть марки Юралс (Urals) [11], сезонно-скорректированный ряд широкой денежной массы (Money) [12], индекс МосБиржи (IMOEX) [13], разница между доходностями долгосрочных (10 лет) и краткосрочных (1 год) облигаций (SpreadIm)[14], стоимость бивалютной корзины (значение, средневзвешенное из курса доллара США с весом 0,55 и Евро с весом 0,45), руб. (Bivalut) [15].

Также были использованы данные мониторинга предприятий, осуществляемого Банком России. Особое внимание уделяется анализу ряда ключевых показателей, которые играют важную роль в оценке текущего состояния экономики и формировании макроэкономических прогнозов на краткосрочную перспективу:

- бизнес-климат Банка России (Бизнес-климат БР) - показатель, отражающий общее состояние и настроения в бизнес-сообществе, основанные на опросах руководителей предприятий;

- объем производства - индикатор, показывающий динамику производственной активности предприятий нефинансового сектора;

- объем производства в прогнозе на 3 месяца - прогнозируемое значение объема производства на ближайшие три месяца, основанное на оценках руководителей предприятий;

- спрос на товары и услуги (индекс спроса) - показатель, отражающий текущую оценку руководством предприятий изменений в спросе на их продукцию;

- спрос на товары и услуги в прогнозе на 3 месяца - ожидаемые изменения в спросе на товары и услуги в течение следующих трех месяцев;

- ценовые ожидания на 3 месяца - прогнозируемые изменения в ценах на отпускные товары

и услуги, ожидаемые респондентами в ближайшие три месяца; цены на готовые товары и услуги - текущий уровень цен на продукцию и услуги, производимые предприятиями;

- издержки производства - стоимостная оценка ресурсов, затрачиваемых на производство товаров и услуг; условия кредитования - оценка руководителями предприятий изменений в условиях получения кредитов[16].

Все данные прошли обработку в Python с помощью библиотеки pandas для повышения их эффективности в обучении моделей машинного обучения. Необходимая предварительная обработка включала удаление лишних столбцов (для каждой модели были выбраны свои предикторы), и нормализация целевого признака (CPI), что позволяет сверточной нейронной сети (CNN) и рекуррентной нейронной сети (LSTM) быстрее обучаться. Также были удалены выбросы, где значение ИПЦ больше 103 - январь 2015 года и март 2022 года (103.85 и 107.61 соответственно).

В тренировочную выборку вошли данные с января 2009 года по декабрь 2022 года включительно, в тестовую вошли данные с января 2023 года по март 2024 года. Создание обучающей и тестовой выборок было важным шагом, позволяющим моделям обучаться на исторических данных и анализировать прогнозы инфляции. Этот подход, учитывающий временной фактор, обеспечивает реалистичное тестирование моделей.

Для улучшения качества прогнозов были созданы сдвинутые версии целевого признака (CPI-1, CPI-2, CPI-3), позволяющие учитывать лаг в прогнозах. Нормализация данных была выполнена с помощью MinMaxScaler из библиотеки scikit-learn.

Разделение данных на обучающую и тестовую выборки является стандартной практикой, позволяющей оценить точность и обобщающую способность модели.

В случае с CNN (свертка нейронных сетей) данные по инфляции были преобразованы в подходящий формат с помощью функции split\_sequence, которая создает последовательность входных данных и соответствующих выходных значений. Это позволяет CNN обучаться на временных последовательностях и выявлять паттерны в данных.

Аналогичным образом, в методе Prophet, созданном Facebook, используются сдвинутые версии целевого признака для учета лага в прогнозах. Этот алгоритм хорошо подходит для временных рядов и может учитывать различные факторы, влияющие на инфляцию.

АНАЛИЗ ПРЕДМЕТНОЙ ОБЛАСТИ

Общие этапы построения моделей в данной работе включали в себя: сбор и подготовка данных, разделение данных, обучение моделей, прогнозирование, оценка эффективности, визуализация результатов, анализ важности признаков, выводы и рекомендации. Для некоторых моделей, таких как градиентный бустинг и случайный лес, проводится также анализ важности признаков, который позволяет понять вклад каждого признака в прогнозирование.

Обучение моделей проводилось с использованием различных поднаборов предикторов. Модели были обучены на тренировочных данных с января 2009 года по декабрь 2022 года включительно и протестированы на тестовых данных с января 2023 года по март 2024 года включительно (15 тестовых точек).

В процессе оптимизации моделей градиентного бустинга и случайного леса для улучшения их эффективности и точности предсказаний были проведены различные эксперименты с разными значениями гиперпараметров. Исходно для градиентного бустинга было выбрано 200 решающих деревьев с максимальной глубиной решающего дерева 3, а для случайного леса 100 решающих деревьев с максимальной глубиной решающего дерева 9, что является распространенным начальным значением для построения данных моделей. Однако в ходе экспериментов и сравнения прогнозов на тестовых данных было принято решение для всех моделей увеличить число деревьев для градиентного бустинга до 500, а для случайного леса до 250. Для моделей, прогнозирующих на 3 месяца вперед была уменьшена глубина решающего дерева для градиентного бустинга до 1, а для случайного леса до 3. Это изменение было мотивировано стремлением улучшить обобщающую способность модели и повысить точность предсказаний. Важность признаков в контексте алгоритмов, основанных на деревьях

решений, может определяться как степень уменьшения неоднородности, которую обеспечивает этот признак. Случайный лес состоит из множества деревьев решений, обученных на случайных подвыборках данных. В каждом дереве определенные признаки используются для разбиений узлов. Для каждого узла рассчитывается уменьшение средней квадратичной ошибки при разбиении данных по данному признаку. Затем суммарное уменьшение средней квадратичной ошибки по всем деревьям нормируется, чтобы определить важность каждого признака. Градиентный бустинг также использует деревья решений, но строит их последовательно, каждая новая модель исправляет ошибки предыдущих. В каждом дереве для каждого признака также рассчитывается уменьшение средней квадратичной ошибки при разбиении узлов. Суммарное уменьшение средней квадратичной ошибки по всем деревьям нормируется, что позволяет определить важность каждого признака. Обученные модели градиентного бустинга и случайного леса затем были протестированы на тестовой выборке, оценивая точность с помощью средней абсолютной ошибки (MAE) [17,18,19].

Признаки, которые ухудшали MAE на тестовых данных отсеивались, начиная с самых неважных. В результате для прогнозирования на 1 и 2 месяца вперед самые точные прогнозы на тестовых данных давали модели случайного леса и градиентного бустинга использующие только значения индекса потребительских цен [20,21,22]. На тестовых данных, при прогнозировании на 3 месяца вперед, лучшие результаты давали модели использующие признаки индекса потребительских цен, среднемесячной фактической ставки по кредитам (в руб.), предоставленным банками (MIACR) и Стоимость бивалютной корзины (значение, средневзвешенное из курса доллара США с весом 0,55 и Евро с весом 0,45), руб. (Bivalut). Для градиентного бустинга важность признаков отражена в таблице 1.

Таблица 1.

Важность признаков для модели градиентного бустинга на 3 месяца вперед  
Table 1. The importance of features for the gradient boosting model for 3 months ahead

Переменные	Важность
Bivalut три месяца назад	0.454
MIACR три месяца назад	0.373
CPI за три месяца назад	0.173

Для случайного леса важность признаков отражена в таблице 2. При построении нейронных сетей были использованы данные только о ИПЦ за три последних месяца. Эти данные были нормализованы с помощью метода MinMaxScaler из библиотеки Scikit-learn. Для прогнозирования

применялись модели рекуррентных нейронных сетей LSTM и сверточных нейронных сетей (CNN). Модели LSTM и сверточные нейронные сети, настроенные с определенными параметрами, продемонстрировали высокую эффективность в задаче прогнозирования инфляции.

Важность признаков для модели случайного леса на 3 месяца вперед  
**Table 2. The importance of features for a random forest model 3 months in advance**

Переменные	Важность
MIACR три месяца назад	0.420
Bivalut три месяца назад	0.404
CPI за три месяца назад	0.175

CNN эффективно захватывают локальные зависимости данных, в то время как LSTM предназначены для работы с долгосрочными зависимостями, что делает их идеальными для задач прогнозирования, таких как прогнозирование инфляции. Эти модели, при правильной настройке и обучении, демонстрируют высокую точность и надёжность в своих прогнозах.

В модели ARIMA значимость коэффициентов определялась на основе P-значений, и в результате был установлен порядок модели (p, d, q) равный (1, 1, 1). Проведенный анализ показал, что коэффициент авторегрессии первого порядка равен 0.675 при стандартной ошибке 0.043. Этот коэффициент статистически значим, что подтверждается P-значением, равным 0.000, Z-статистикой 15.516 и 95% доверительным интервалом от 0.590 до 0.760. Коэффициент скользящего среднего первого порядка равен -0.982 при стандартной ошибке 0.036 и также является статистически значимым. Это подтверждается P-значением, равным 0.000, Z-статистикой -27.528 и 95% доверительным интервалом от -1.052 до -0.912. Оценка дисперсии ошибки составляет 0.131 при стандартной ошибке 0.009 и является статистически значимой, что подтверждается P-значением, равным 0.000, Z-статистикой 14.054 и 95% доверительным интервалом от 0.113 до 0.150.

Модель характеризуется следующими критериями информационного отбора: критерий информационного отбора (AIC) составляет 140.632; Байеса (BIC) - 149.931, Ханнана-Куинна (HQIC) - 144.407. Диагностика модели включала Ljung-Box тест на автокорреляцию остатков, где значение статистики Q составляет 1.05 при P-значении 0.30, что указывает на отсутствие значимой автокорреляции остатков. Тем не менее, Jarque-Bera тест на нормальность распределения остатков показал значение статистики JB 99.56 при P-значении 0.00, что свидетельствует о ненормальном распределении остатков. Тест на гетероскедастичность показал значение статистики H 0.47 при P-значении 0.01, указывая на наличие гетероскедастичности. Коэффициент асимметрии остатков равен 0.52, а коэффициент куртозиса - 6.67, что указывает на положительную асиммет-

рию и выраженную куртозис.

На основании полученных результатов можно заключить, что модель ARIMA (1,1,1) демонстрирует высокую значимость включенных компонент, что подтверждается низкими P-значениями для всех коэффициентов. Однако несмотря на отсутствие значимой автокорреляции остатков, тесты на нормальность распределения и гетероскедастичность указывают на определенные проблемы, которые могут повлиять на надежность выводов модели. В частности, ненормальность распределения остатков и наличие гетероскедастичности требуют дополнительного внимания и, возможно, применения дополнительных методов моделирования или преобразования данных. В целом, полученные результаты подтверждают адекватность модели ARIMA (1,1,1) для анализа временного ряда ИПЦ, но также указывают на направления для дальнейшего улучшения модели.

Модель векторной авторегрессии (VAR) для анализа взаимосвязи между индексом потребительских цен и спросом на товары и услуги на три месяца вперед была обучена методом наименьших квадратов (OLS). Эта модель включает два уравнения. Критерий информационного отбора (AIC) составил 0.262, Байеса (BIC) - 0.375, Ханнана-Куинна (HQIC) - 0.308. Логарифм правдоподобия модели равен -480.901, а финальная прогнозная ошибка (FPE) составила 1.300.

Рассматривая результаты для уравнения ИПЦ, константа имеет коэффициент 37.240 при стандартной ошибке 5.659, t-статистика равна 6.581, а P-значение 0.000, что указывает на статистическую значимость. Лаг ИПЦ имеет коэффициент 0.628 при стандартной ошибке 0.056, t-статистика составляет 11.167, а P-значение равно 0.000, подтверждая его значимость. Лаг спроса на товары и услуги за три месяца имеет коэффициент 0.012 при стандартной ошибке 0.005, t-статистика 2.525 и P-значение 0.012, что также указывает на значимость.

Для уравнения спроса на товары и услуги за три месяца константа имеет коэффициент 67.180 при стандартной ошибке 55.522, t-статистика равна 1.210 и P-значение 0.226, что указывает на его незначимость. Лаг ИПЦ имеет

коэффициент  $-0.645$  при стандартной ошибке  $0.552$ ,  $t$ -статистика равна  $-1.169$  и  $P$ -значение  $0.242$ , также не указывая на значимость. Однако лаг спроса на товары и услуги за три месяца имеет коэффициент  $0.780$  при стандартной ошибке  $0.048$ ,  $t$ -статистика  $16.256$  и  $P$ -значение  $0.000$ , что подтверждает его значимость.

Корреляционная матрица остатков показала, что корреляция между остатками ИПЦ и спросом на товары и услуги за три месяца равна  $-0.180714$ , что указывает на слабую отрицательную корреляцию.

Таким образом, модель VAR демонстрирует, что ИПЦ значимо зависит от своего значения за прошлый период и значения спроса на товары и услуги за прошлый период. Спрос на товары и услуги за три месяца значимо зависит только от собственного лага, что указывает на автокорреляцию и устойчивость этой переменной во времени. Эти выводы подтверждают адекватность модели VAR для анализа взаимосвязи между индексом потребительских цен и спросом на товары и услуги за три месяца вперед и их прогнозирования на один, два и три месяца. В модели Prophet использовались данные индекса потребительских цен, которые были преобразованы с добавлением лагов, и эти значения подавались модели в качестве

регрессора. Также модель использует месячную сезонную компоненту. Дополнительно модель Prophet включает в себя тренд, который описывает долгосрочное направление данных. Эти компоненты позволяют более точно моделировать и предсказывать временные ряды, учитывая, как краткосрочные колебания, так и долгосрочные тенденции.

## РЕЗУЛЬТАТЫ И ДИСКУССИИ

Для оценки точности модели, показывающей, насколько предсказания модели близки к фактическим значениям в данной работе применяется метрика MAE, где показание вычисляется как среднее абсолютных разностей между прогнозами и реальными значениями. Чем ниже показатель метрики, тем лучше модель соответствует набору данных [17].

Результаты градиентного бустинга показали высокую точность прогнозирования инфляции на тестовых данных для 1 и 2 месяцев вперед, но для 3 месяцев вперед модель начинает завышать инфляцию, что приводит к большим ошибкам на тестовой выборке. Ниже представлен рисунок 1 на котором изображены графики с прогнозами на 1, 2 и 3 месяца вперед.

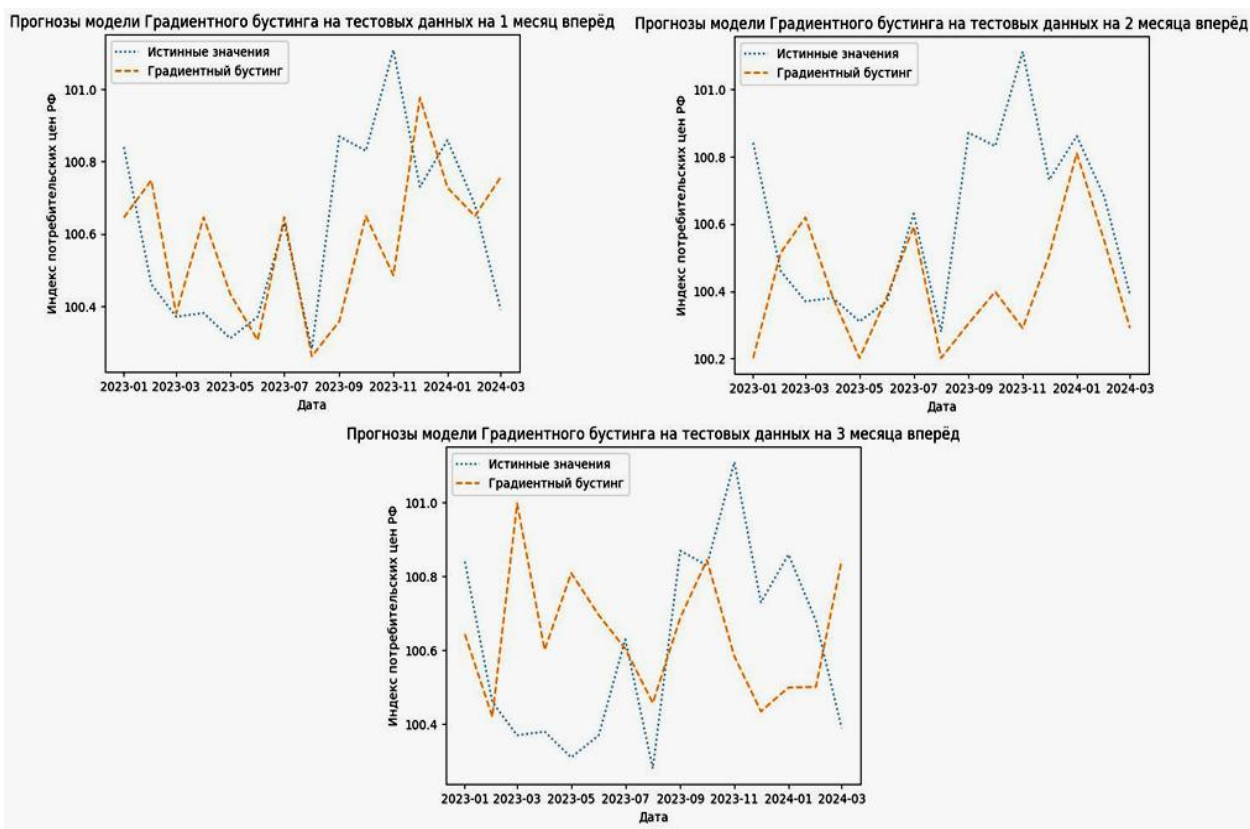


Рисунок 1. Прогнозы модели Градиентного бустинга на тестовых данных на 1, 2 и 3 месяца вперед  
 Figure 1. Gradient boosting model forecasts based on test data for 1, 2 and 3 months ahead

Средняя абсолютная ошибка для модели градиентного бустинга, прогнозирующей на 1 месяц вперед, составила 0.205, для модели, прогнозирующей на 2 месяца вперед, составила 0.233, а для модели, прогнозирующей на 3 месяца вперед, ошибка составила 0.275. Результаты случайного леса показали высокую точность прогнозирования инфляции на тестовых данных для 1 месяца впе-

ред, для прогнозирования на 2 месяца вперед модель показала самый лучший результат по метрике MAE, но для 3 месяцев вперед модель начинает завышать инфляцию, что приводит к большим ошибкам на тестовой выборке. Ниже представлен рисунок 2 на котором изображены графики с прогнозами на 1, 2 и 3 месяца вперед.

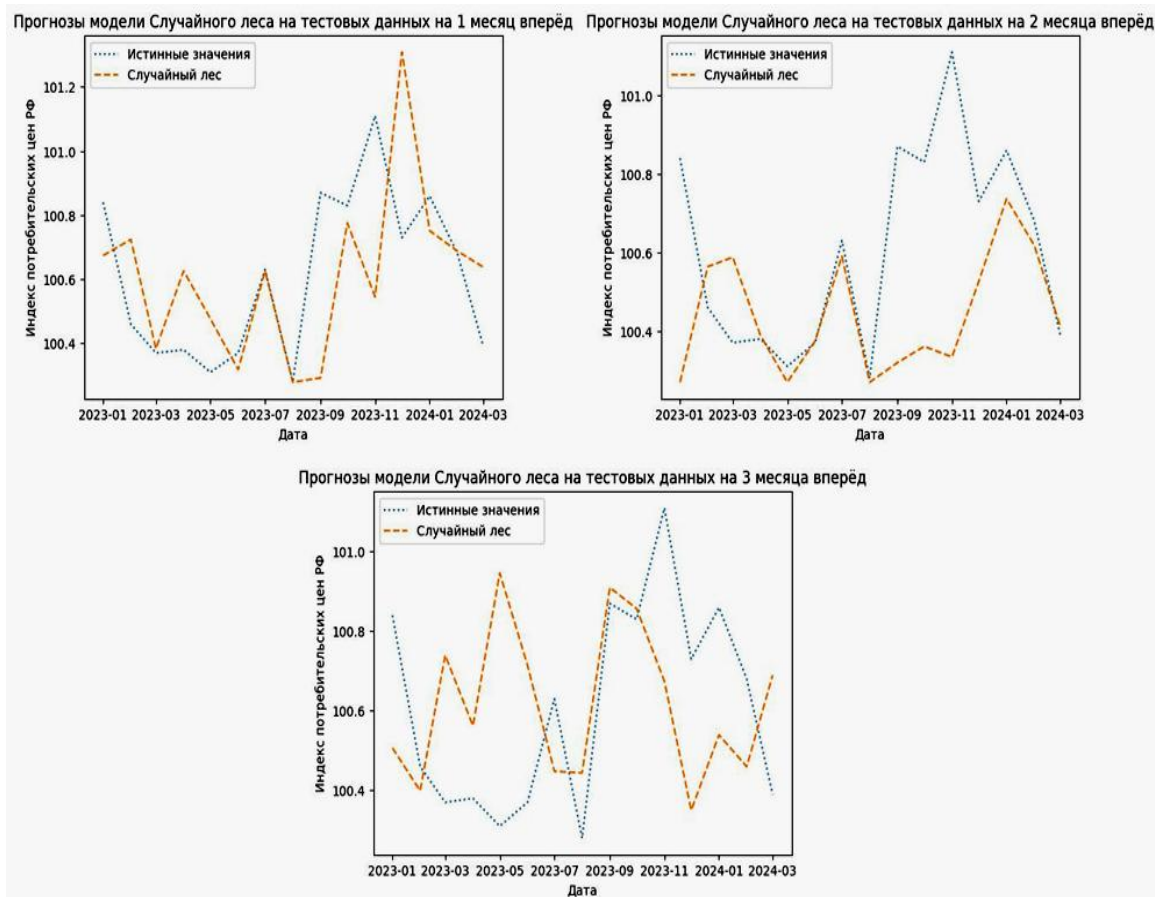


Рисунок 2. Прогнозы модели случайного леса на тестовых данных на 1, 2 и 3 месяца вперед  
 Figure 2. Forecasts of the random forest model based on test data for 1, 2 and 3 months ahead

Средняя абсолютная ошибка для модели случайного леса, прогнозирующей на 1 месяц вперед, составила 0.203, для модели, прогнозирующей на 2 месяца вперед, показала самый лучший результат по метрике MAE, составила 0.213, а для модели, прогнозирующей на 3 месяца вперед, ошибка составила 0.266. Результаты LSTM модели показали высокую точность прогнозирования инфляции на тестовых данных для 1, 2 и 3 месяца вперед. Ниже представлен рисунок 3 на котором изображены графики с прогнозами на 1, 2 и 3 месяца вперед. Средняя абсолютная ошибка для LSTM модели, прогнозирующей на 1 месяц вперед, составила 0.192, для модели, прогнозирующей на 2 месяца вперед, составила 0.216, а для

модели, прогнозирующей на 3 месяца вперед, ошибка составила 0.239. Результаты CNN модели показали высокую точность прогнозирования инфляции на тестовых данных для 1 и 2 месяцев вперед, для прогнозирования на 3 месяца вперед модель показала самый лучший результат по метрике MAE. Ниже представлен рисунок 4 на котором изображены графики с прогнозами на 1, 2 и 3 месяца вперед.

Средняя абсолютная ошибка для CNN модели, прогнозирующей на 1 месяц вперед, составила 0.186, для модели, прогнозирующей на 2 месяца вперед, составила 0.223, а для модели, прогнозирующей на 3 месяца вперед, ошибка оказалась составила 0.225.

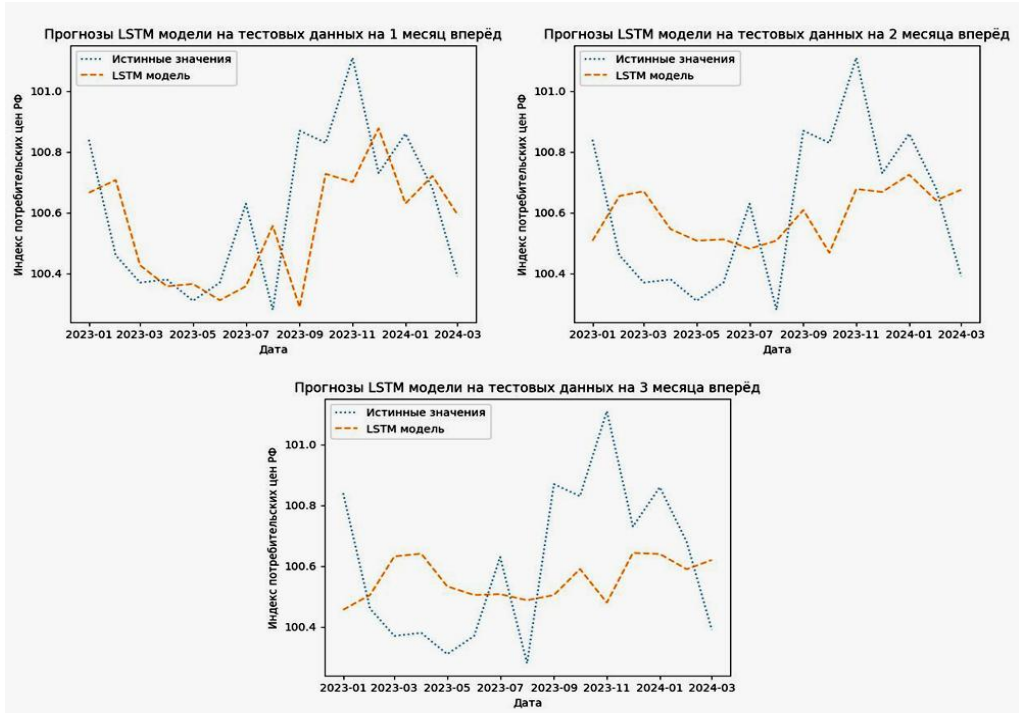


Рисунок 3. Прогнозы модели LSTM на тестовых данных на 1, 2 и 3 месяца вперед  
Figure 3. Forecasts of the LSTM model based on test data for 1, 2 and 3 months ahead

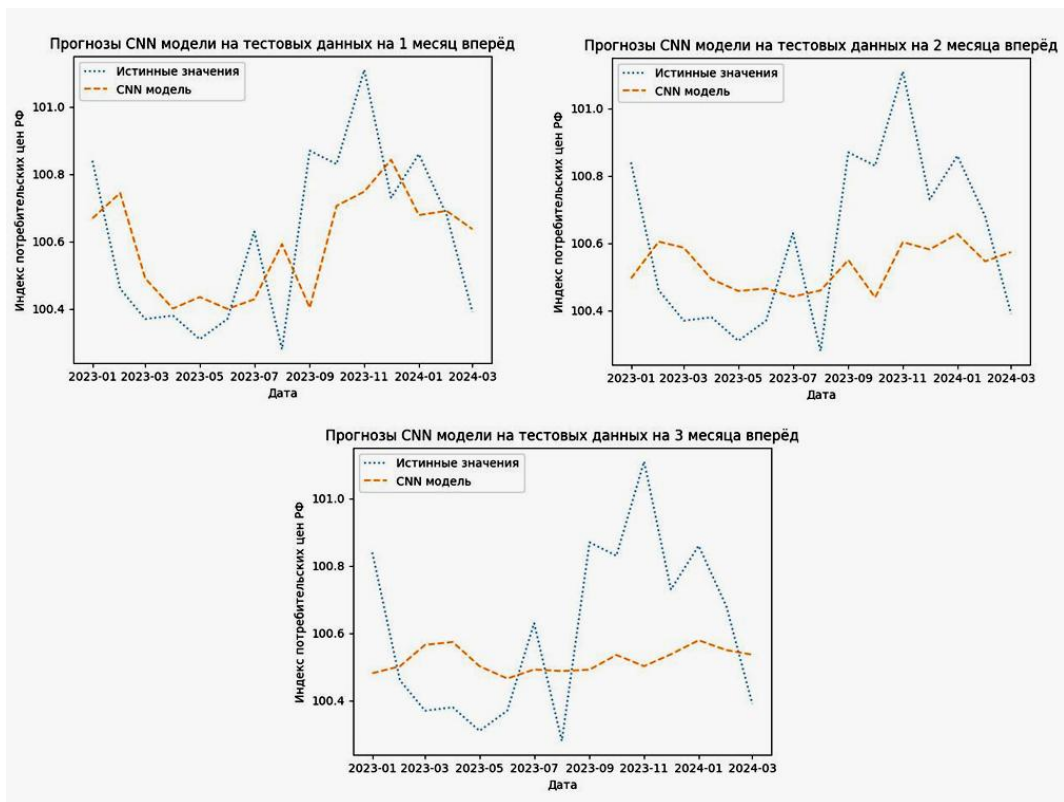


Рисунок 4. Прогнозы CNN модели на тестовых данных на 1, 2 и 3 месяца вперед  
Figure 4. CNN model forecasts based on test data for 1, 2 and 3 months ahead



Средняя абсолютная ошибка для CNN модели, прогнозирующей на 1 месяц вперед, составила 0.186, для модели, прогнозирующей на 2 месяца вперед, составила 0.223, а для модели, прогнозирующей на 3 месяца вперед, ошибка оказалась составила 0.225.

Результаты ARIMA модели показали высокую точность прогнозирования инфляции на тестовых данных для 1, 2 и 3 месяцев вперед. Ниже представлен рисунок 5 на котором изображены графики с прогнозами на 1, 2 и 3 месяца вперед.

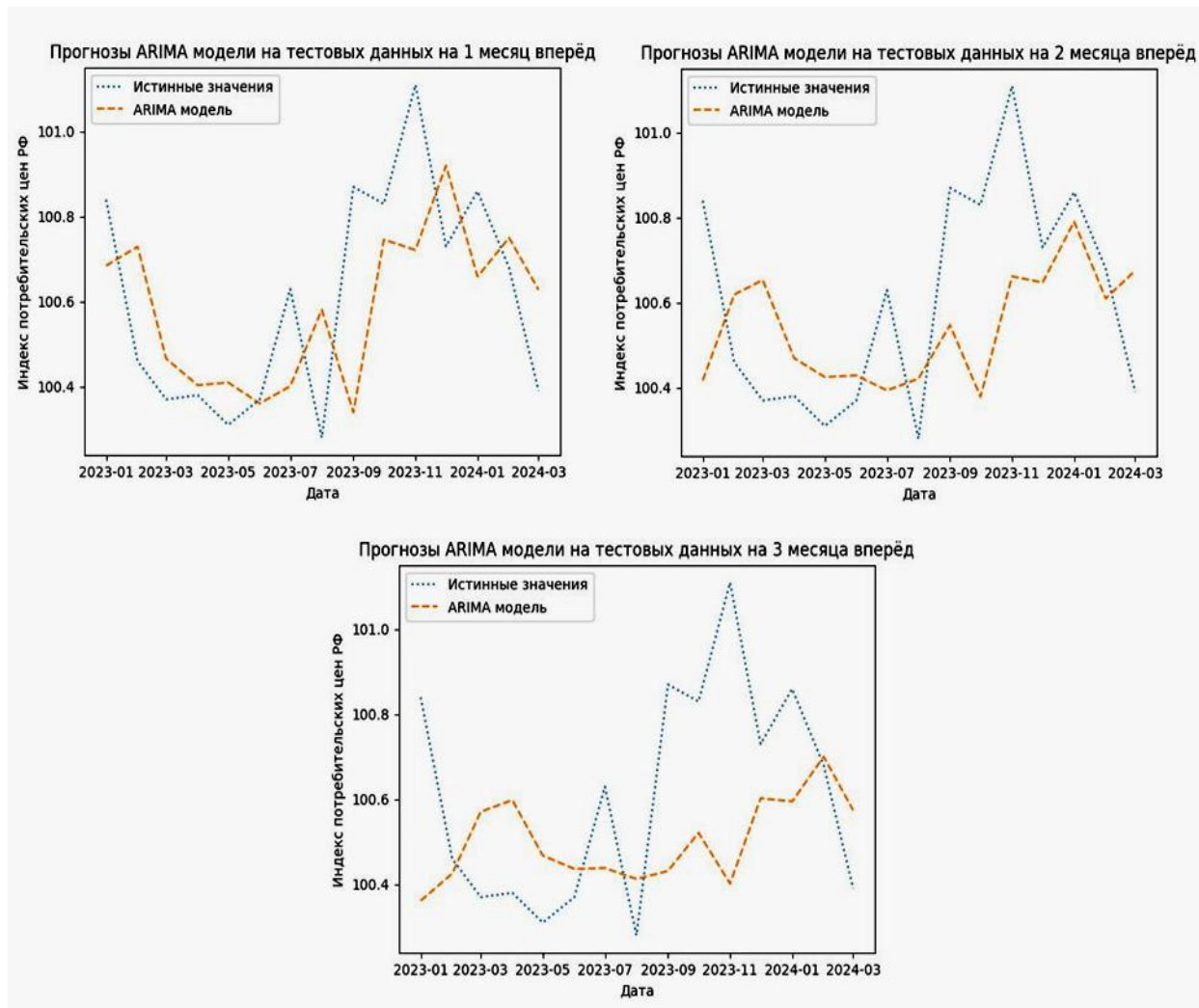


Рисунок 5. Прогнозы ARIMA модели на тестовых данных на 1, 2 и 3 месяца вперед  
Figure 5. ARIMA model forecasts based on test data for 1, 2 and 3 months ahead

Средняя абсолютная ошибка для ARIMA модели, прогнозирующей на 1 месяц вперед, составила 0.192, для модели, прогнозирующей на 2 месяца вперед, составила 0.216, а для модели, прогнозирующей на 3 месяца вперед, ошибка составила 0.236. Результаты VAR модели показали высокую точность прогнозирования инфляции на тестовых данных для 1, 2 и 3 месяцев вперед. Ниже представлен рисунок 6 на котором изображены графики с прогнозами на 1, 2 и 3 месяца вперед.

Средняя абсолютная ошибка для VAR модели, прогнозирующей на 1 месяц вперед, составила 0.204, для модели, прогнозирующей на 2 месяца вперед, составила 0.218, а для модели, прогнозирующей

на 3 месяца вперед, ошибка составила 0.238. Результаты Prophet модели показали высокую точность прогнозирования инфляции на тестовых данных для 1, 2 и 3 месяцев вперед, для прогнозирования на 1 месяц вперед модель показала самый лучший результат по метрике MAE. Ниже представлен рисунок 7 на котором изображены графики с прогнозами на 1, 2 и 3 месяца вперед.

Средняя абсолютная ошибка для Prophet модели, прогнозирующей на 1 месяц вперед, составила 0.174, для модели, прогнозирующей на 2 месяца вперед, составила 0.219, а для модели, прогнозирующей на 3 месяца вперед, ошибка составила 0.227.

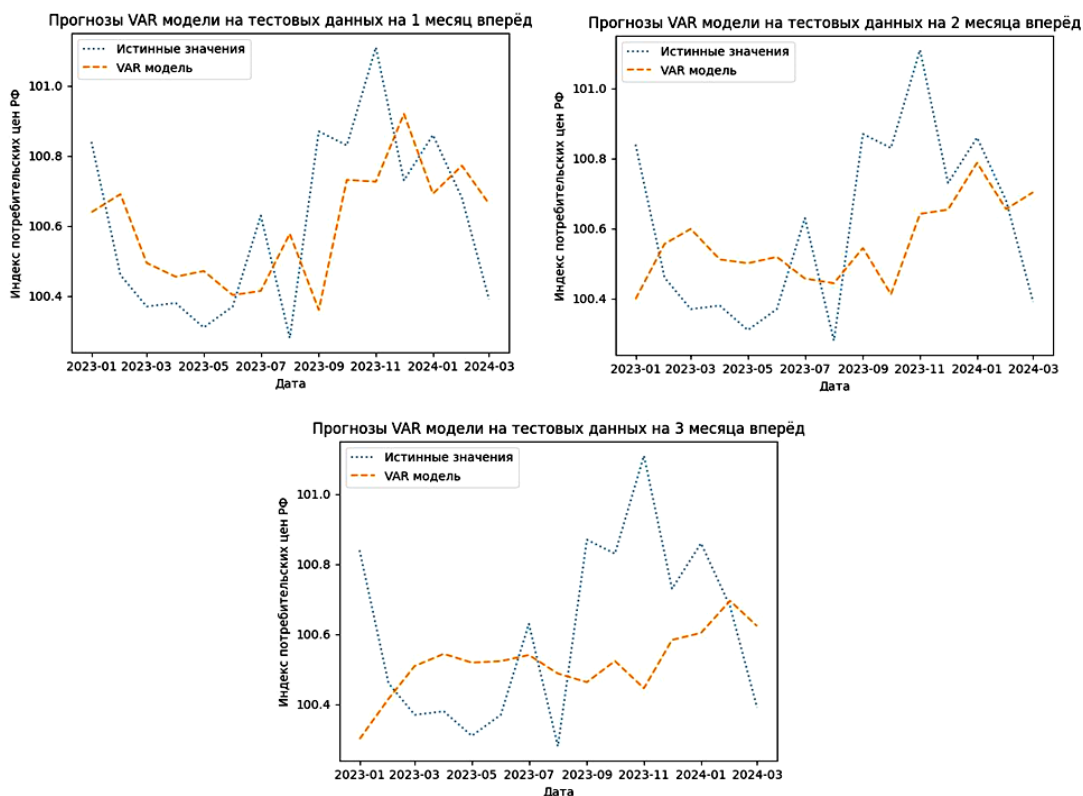


Рисунок 6. Прогнозы VAR модели на тестовых данных на 1, 2 и 3 месяца вперед  
 Figure 6. VAR model forecasts based on test data for 1, 2 and 3 months ahead

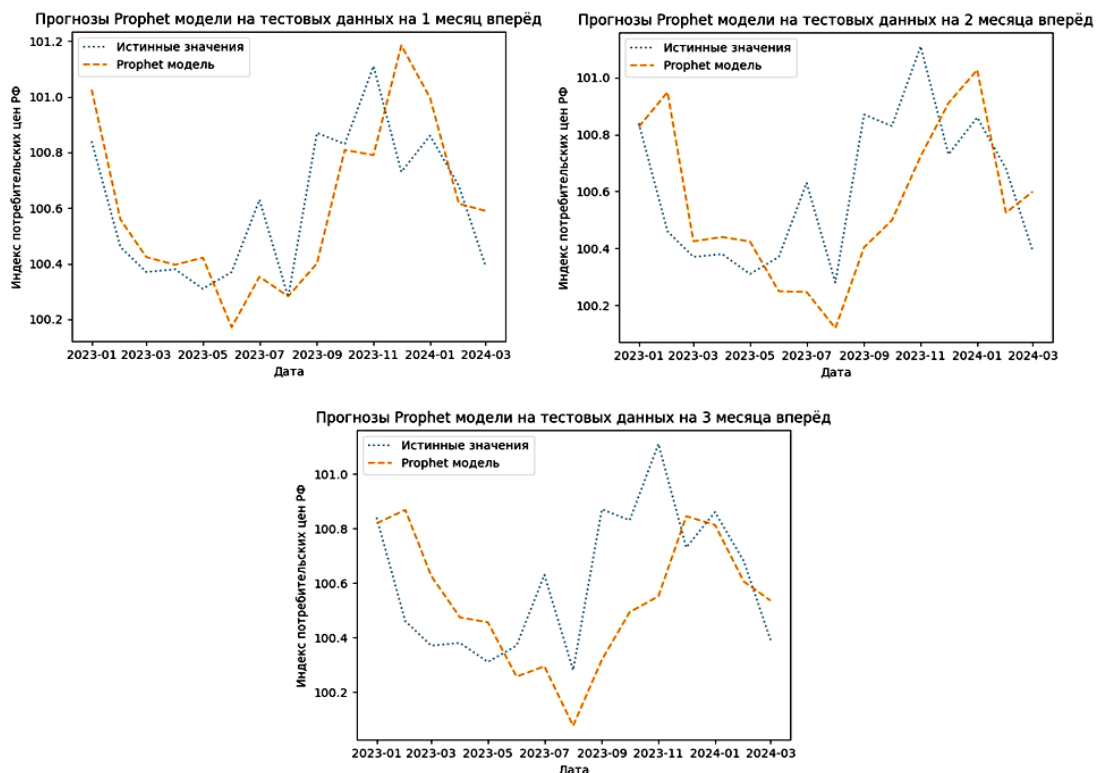


Рисунок 7. Прогнозы Prophet модели на тестовых данных на 1, 2 и 3 месяца вперед  
 Figure 7. The forecasts of the Prophet model based on test data for 1, 2 and 3 months ahead

На основе анализа результатов мы можем сделать вывод, что для прогнозирования CPI на 1 месяц вперед лучше всего подходят Prophet, CNN и LSTM. Для прогнозирования на 2 месяца вперед лучше всего подходят случайный лес, ARIMA и LSTM. Для прогнозирования на 3 месяца вперед лучше всего подходят CNN, Prophet и ARIMA. Выбор модели зависит от конкретных требований

и доступных данных. Ниже приводится таблица 3 всех семи моделей с рассчитанной MAE метрикой на тестовых данных при прогнозировании на 1, 2 и 3 месяца вперед и их сравнение с наивной моделью. Все модели сравниваются с наивной моделью, которая в качестве прогноза берёт значение индекса потребительских цен за последний месяц.

Таблица 3.

Значение MAE для 7 моделей и наивной модели на тестовых данных при прогнозировании на 1, 2 и 3 месяца вперед

Table 3. The value of MAE for 7 models and a naive model based on test data when forecasting for 1, 2 and 3 months ahead

Модель	MAE на 1 месяц	MAE на 2 месяца	MAE на 3 месяца
Prophet	0.174	0.219	0.227
CNN	0.186	0.223	0.225
LSTM	0.192	0.216	0.239
ARIMA	0.192	0.216	0.236
VAR	0.204	0.218	0.238
Случайный лес	0.203	0.213	0.266
Градиентный бустинг	0.205	0.234	0.275
Наивная модель	0.211	0.248	0.301

### ЗАКЛЮЧЕНИЕ

В ходе анализа результатов прогнозирования инфляции с использованием различных моделей было выявлено, что каждая модель имеет свои сильные и слабые стороны в зависимости от горизонта прогнозирования. Модели градиентного бустинга и случайного леса продемонстрировали высокую точность при прогнозировании инфляции на 1 и 2 месяца вперед, где модель случайного леса показала наиболее точные прогнозы для 2 месяцев среди всех моделей, однако на 3 месяца вперед их прогнозы были менее точными, что привело к увеличению средней абсолютной ошибки (MAE). В частности, MAE для градиентного бустинга составляла 0.205, 0.233 и 0.275 для прогнозов на 1, 2 и 3 месяца соответственно, в то время как для случайного леса эти значения были 0.203, 0.213 и 0.266.

Среди более сложных моделей машинного обучения, LSTM и CNN также показали хорошие результаты. Модель LSTM продемонстрировала высокую точность на всех горизонтах прогнозирования, с MAE 0.192, 0.216 и 0.239 для 1, 2 и 3 месяцев вперед соответственно. CNN модель показала хороший результат для прогноза на 1 и 2 месяца вперед с MAE 0.186 и 0.223 соответственно, но также оказалась наиболее точной для 3 месяцев вперед с MAE 0.225, что является наилучшим показателем среди всех моделей на этом горизонте. Традиционные временные ряды, такие как ARIMA и VAR, также продемонстрировали

высокую точность прогнозирования. ARIMA показала стабильные результаты с MAE 0.192, 0.216 и 0.236 для 1, 2 и 3 месяцев вперед соответственно, что очень близко к показателям LSTM. VAR модель продемонстрировала аналогичную стабильность, с MAE 0.204, 0.218 и 0.238.

Наиболее качественные результаты были получены с использованием модели Prophet. Prophet показала наилучший результат для прогноза на 1 месяц вперед с MAE 0.174, что является наименьшим значением среди всех моделей. Для прогнозов на 2 и 3 месяца вперед MAE составила 0.219 и 0.227 соответственно, что также является конкурентоспособным показателем по сравнению с другими моделями.

Результаты показали, что выбор модели для прогнозирования инфляции зависит от требуемого горизонта прогнозирования. Для прогнозов на 1 месяц вперед модель Prophet оказалась наиболее точной, в то время как на 2 месяца вперед лучшие результаты показала модель случайного леса, а модель CNN показала лучшие результаты на 3 месяца вперед.

Успешное прогнозирование требует сочетания эффективных моделей, качественных данных и адаптации к меняющимся условиям. Непрерывное обучение, тестирование и улучшение моделей могут помочь повысить точность прогнозов и обеспечить более информированные решения. Развитие технологий и доступ к большим данным открывают новые направления в прогнозировании инфляции.

ЛИТЕРАТУРА

REFERENECES

1. Астраханцева И.А., Герасимов А.С. Прогнозирование региональной инфляции на основе гибридной модели машинного обучения: градиентный бустинг и случайный лес НАУЧНЫЕ ТРУДЫ Вольного экономического общества России – 2023. № 5. С. 200-226. DOI: 10.38197/2072-2060-2023-243-5-200-226
2. Астраханцева И.А., Герасимов А.С., Астраханцев Р.Г. Прогнозирование региональной инфляции с помощью алгоритмов машинного обучения. *Известия высших учебных заведений. Серия «Экономика, финансы и управление производством» [Ивэкофин]*. 2022. № 4(54). С. 6–13. DOI 10.6060/ivecofin.2022544.620.
3. Банк России. Что такое инфляция.2020. / www.cbr.ru.
4. Косарев А.Е., Козарезова Л.О. Индексы цен // Большая российская энциклопедия: [в 35 т.] / гл. ред. Ю. С. Осипов. — М.: Большая российская энциклопедия
5. Таблицы инфляции. <https://уровень-инфляции.рф> Официальный сайт Федеральной службы государственной статистики. <http://www.gks.ru>.
6. Методология расчета индексов потребительских цен 2008 // Федеральная служба государственной статистики.
7. Павлов Е. Прогнозирование инфляции в России с помощью нейронных сетей // Деньги и кредит. – март 2020. – т.79 № 1. – pp. 57-73. – doi: 10.31477/rjmf.202001.57
8. Балацкий Е.В., Юревич М.А. Прогнозирование инфляции: практика использования синтетических процедур. *Мир новой экономики*. 2018. №4. С.20–31.
9. Показатели ставок межбанковского кредитного рынка группы МIACR с 01.08.2000 [http://www.cbr.ru/hd\\_base](http://www.cbr.ru/hd_base).
10. Спотовая цена сырой нефти Urals: цена и курс.
11. Банк России. Денежно-кредитная и финансовая статистика.[https://www.cbr.ru/statistics/macro\\_itm/dkfs](https://www.cbr.ru/statistics/macro_itm/dkfs)
12. Индекс МосБиржи. <https://www.moex.com/ru>
13. Спред между доходностями долгосрочных (10 лет) и среднесрочных (1 год) облигаций. [http://www.cbr.ru/hd\\_base](http://www.cbr.ru/hd_base).
14. Банк России. Мониторинг предприятий 2024 г. /<https://cbr.ru/dkp/mp/> (дата обращения: 20.05.2024).
15. Уилмотт, Корт Дж.; Матсуура, Кэндзи (19 декабря 2005 г.). "Преимущества средней абсолютной ошибки (MAE) перед среднеквадратичной ошибкой (RMSE) при оценке средних характеристик модели". *Климатические исследования*. 30: 79-82. doi: 10.3354/cr030079.
16. Астраханцева, И. А. Системный подход к анализу фрактальной природы сложных технических систем. *Известия высших учебных заведений. Серия «Экономика, финансы и управление производством» [Ивэкофин]*. 2023. № 3(57). С. 89-97. DOI 10.6060/ivecofin.2023573.657.
17. Горев, С. В. Исследование методов и алгоритмов искусственного интеллекта при определении стоимости произведений искусства. *Известия высших учебных заведений. Серия «Экономика, финансы и управление производством» [Ивэкофин]*. 2022. – № 4(54). – С. 21-28. – DOI 10.6060/ivecofin.2022544.622.
18. Морозов, Е. Н. Математические модели для оптимизации машинночитаемых регулятивных систем. *Известия высших учебных заведений. Серия «Экономика, финансы и управление производством» [Ивэкофин]*. 2023. № 4(58). С. 71-78. DOI 10.6060/ivecofin.2023584.666.
19. Астраханцева, И. А. Фрактальный анализ в оценке эффективности и надежности сложных технических систем. *Современные наукоемкие технологии. Региональное приложение*. 2023. № 4(76). С. 60-68. DOI 10.6060/snt.20237604.0008. – EDN NBDYHR.
20. Горев, С. В. Анализ практики налогообложения сделок с предметами искусства. *Известия высших учебных заведений. Серия «Экономика, финансы и управление производством» [Ивэкофин]*. 2021. № 2(48). С. 23-28. – DOI 10.6060/ivecofin.2021482.530.
1. Astrakhantseva I.A., Gerasimov A.S. Forecasting regional inflation based on a hybrid machine learning model: gradient boosting and random forest SCIENTIFIC PAPERS of the Free Economic Society of Russia – 2023. – No. 5. – pp. 200-226. – DOI: 10.38197/2072-2060-2023-243-5-200-226
2. Astrakhantseva I.A., Gerasimov A.S., Astrakhantsev R.G. Forecasting regional inflation using machine learning algorithms. *Ivecofin*. 2022. № 4(54). P. 6-13. DOI 10.6060/ivecofin.2022544.620.
3. The Bank of Russia. What is inflation?2020. / www.cbr.ru.
4. Kosarev A.E., Kozarezova L.O. Price indices // The Great Russian Encyclopedia: [in 35 volumes] / ch. ed. Yu. S. Osipov. — M.: The Great Russian Encyclopedia
5. Tables of inflation. <https://level-of-inflation.Russian Federation> The official website of the Federal State Statistics Service. <http://www.gks.ru>.
6. Methodology for calculating consumer price indices 2008 // Federal State Statistics Service. / [http://www.gks.ru/free\\_doc/new\\_site/prices/ipc\\_met.htm](http://www.gks.ru/free_doc/new_site/prices/ipc_met.htm)
7. Pavlov E. Forecasting inflation in Russia using neural networks // Money and credit. – March 2020. – vol.79 No. 1. – pp. 57-73. – doi: 10.31477/rjmf.202001.57.
8. Balatsky E.V., Yurevich M.A. Forecasting inflation: the practice of using synthetic procedures. *The world of the new economy*. 2018. No.4. pp.20-31.
9. Rates of the interbank credit market of the MIACR Group from 08/01/2000 /[http://www.cbr.ru/hd\\_base/mkr/mkr\\_base](http://www.cbr.ru/hd_base/mkr/mkr_base)
10. Spot price of Urals crude oil: price and exchange rate <https://ru.investing.com/commodities/crude-oil-urals-spot-futures>
11. The Bank of Russia. Monetary and financial statistics. [https://www.cbr.ru/statistics/macro\\_itm/dkfs](https://www.cbr.ru/statistics/macro_itm/dkfs)
12. Moscow Exchange Index /<https://www.moex.com/ru>
13. The spread between the yields of long-term (10 years) and medium-term (1 year) bonds./[http://www.cbr.ru/hd\\_base](http://www.cbr.ru/hd_base).
14. The dual currency basket of the Central Bank of the Russian Federation /[https://kurs.vip/currencies/cbr/bivalutnaya\\_korzina](https://kurs.vip/currencies/cbr/bivalutnaya_korzina) The Bank of Russia. Monitoring of enterprises in 2024 /<https://cbr.ru/dkp/mp/>
15. Wilmott, Kort J.; Matsuura, Kenji (December 19, 2005). "Advantages of the mean absolute error (MAE) over the mean square error (RMSE) in evaluating the average characteristics of the model." *Climate research*. 30: 79-82. doi: 10.3354/cr030079.
16. Astrakhantseva I. A. Systematic approach to the analysis of the fractal nature of complex technical systems *Ivecofin*. 2023. No. 3(57). P. 89-97. DOI 10.6060/ivecofin.2023573.657.
17. Gorev, S. V. Research methods and algorithms of artificial intelligence in determining the value of discovered art. *Ivecofin*. 2022. No. 4(54). P. 21-28. DOI 10.6060/ivecofin.2022544.622.
18. Morozov, E. N. Mathematical models for optimization of machine-readable regulatory systems. *Ivecofin*. 2023. No. 4(58). – pp. 71-78. – DOI 10.6060/ivecofin. 2023-584.666.
19. Astrakhantseva I. A. Fractal analysis of the efficiency and stability of complex technical systems / I. A. Astrakhantseva, S. V. Gorev, R. G. Astrakhantsev // *Modern science-intensive technologies. Regional application*. – 2023. – No. 4(76). – P. 60-68. – DOI 10.6060/snt.20237604.0008. – EDN NBDYR.
20. Gorev, S. V. Analysis of the practice of taxation of transactions with objects of art. *Ivecofin*. 2021. – No. 2(48). – pp. 23-28. – DOI 10.6060/ivecofin.2021482.530.

Поступила в редакцию (Received) 10.05.2024  
Принята к опубликованию (Accepted) 10.07.2024