

# Прогнозирование инфляции в России на основе градиентного бустинга и нейронных сетей<sup>1</sup>

Урмат Джункеев, РАНХиГС; МГУ

dzhunkeev@gmail.com

Цель работы – оценка качества прогнозирования инфляции в России при помощи методов машинного обучения – алгоритмов градиентного бустинга и нейронных сетей. Это первая работа, в которой для прогнозирования инфляции в России применяются модели долгой краткосрочной памяти и управляемого рекуррентного блока. Кроме того, мы апробируем модифицированные версии градиентного бустинга – LightGBM и CatBoost. На выборке из лаговых величин инфляции наиболее точные прогнозы получены с помощью сверточных и полносвязных нейронных сетей, а при прогнозировании на горизонте в 12 месяцев – при использовании модели долгой краткосрочной памяти, что связано с последовательной обработкой информации и пропускным механизмом в изучении статистических данных. При учете дополнительных макроэкономических факторов прогнозное преимущество демонстрируют полносвязные нейронные сети и модель градиентного бустинга Sklearn. Согласно декомпозиции Шепли, наиболее информативными предикторами при прогнозировании российской инфляции являются цены на нефть и природный газ, инфляция в еврозоне и США, динамика оборота розничной торговли, темп прироста численности безработных.

## Ключевые слова:

*прогноз инфляции, машинное обучение, градиентный бустинг, нейронные сети, вектор Шепли*

**JEL Codes:** C52, C53, E37

## Цитирование:

Dzhunkeev, U. (2024). Forecasting Inflation in Russia Using Gradient Boosting and Neural Networks. *Russian Journal of Money and Finance*, 83(1), pp. 53–76.

## 1. Введение

С переходом Банка России на таргетирование инфляции задача ее прогнозирования представляется все более актуальной. Важность повышения точности прогнозирования инфляции объясняется тем, что решения по денежно-кредитной политике опираются не только на текущую информацию об экономической ситуации, но и на прогнозы будущей макроэкономической конъюнктуры (Svensson, 2010).

<sup>1</sup> Автор выражает благодарность трем анонимным рецензентам, Александре Божечковой (ИЭП им. Е. Т. Гайдара; РАНХиГС), Юрию Перевышину (РАНХиГС), Андрею Полбину (ИЭП им. Е. Т. Гайдара; РАНХиГС) и участникам XXIV Ясинской (Апрельской) конференции ВШЭ 2023 г. за ценные комментарии, которые позволили улучшить работу.

Помимо этого краткосрочные прогнозы используются Банком России для формирования предпосылок среднесрочного прогноза, исходя из которого принимаются решения по денежно-кредитной политике (Андреев, 2016). Наряду с центральным банком в получении более точных прогнозов инфляции заинтересованы домашние хозяйства и частные фирмы вследствие того, что трудовые контракты, долговые обязательства и продажи устанавливаются в номинальном выражении (Medeiros et al., 2019). Неопределенность будущей динамики инфляции сопровождается ростом процентных ставок, что может привести к снижению уровня благосостояния и экономической активности (Кудрин, 2007).

В литературе отмечается, что превзойти точность прогнозов инфляции, полученных на основе эталонных моделей, например моделей случайного блуждания (Random Walk, RW), моделей ненаблюдаемых компонент со стохастической волатильностью (Unobserved Components with Stochastic Volatility, UCSV) или авторегрессий (Autoregression, AR) первого порядка, – трудновыполнимая задача (Atkenson and Ohanian, 2001; Stock and Watson, 2007; Faust and Wright, 2013; Joseph et al., 2021). Однако ряд авторов показывают, что применение методов машинного обучения позволяет решить данную задачу с меньшей ошибкой прогноза в сравнении с более простыми бенчмарками (Chakraborty and Joseph, 2017; Garcia et al., 2017; Medeiros et al., 2019; Paranhos, 2021; Almosova and Andresen, 2023; Araujo and Gaglianone, 2023). Этот вывод эмпирически подтверждают ряд российских исследователей (Baybuza, 2018; Pavlov, 2020; Shulyak, 2022; Semiturkin and Shevelev, 2023).

В отличие от Baybuza (2018), Medeiros et al. (2019), Mamedli and Shibitov (2021) и Shulyak (2022) мы применяем более продвинутые алгоритмы градиентного бустинга (gradient boosting): LightGBM (Ke et al., 2017) и CatBoost (Dorogush et al., 2018; Prokhorenkova et al., 2018). Также мы дополняем исследования Nakamura (2005), Szafranek (2019), Pavlov (2020) и Joseph et al. (2021), используя для прогнозирования сверточные (Convolutional Neural Networks, CNN) и рекуррентные нейронные сети (Recurrent Neural Networks, RNN). Кроме того, мы совершенствуем подход Paranhos (2021) и Almosova and Andresen (2023) к применению моделей долгой краткосрочной памяти (Long Short-Term Memory, LSTM), проводя обучение модели управляемого рекуррентного блока (Gated Recurrent Unit, GRU) и многофакторных RNN соответственно. Наконец, помимо моделей GRU, которые применяют Barkan et al. (2023), мы оцениваем CNN.

Далее работа структурирована следующим образом. Раздел 2 содержит обзор литературы, в котором представлены результаты применения методов машинного обучения при прогнозировании инфляции. В Разделе 3 изложена методология исследования. В Разделе 4 дано описание выборочной совокупности. В Разделе 5 приводятся результаты прогнозирования инфляции. В Разделе 6 сформулированы выводы и направления дальнейших исследований.

## 2. Обзор литературы

В данном разделе описываются научные исследования, в которых применялись эконометрические методы и методы машинного обучения для прогнозирования инфляции. Методы прогнозирования инфляции мы поделили на четыре группы: *эконометрические, байесовские, ансамблевые, нейросетевые*. В конце раздела

перечисляются ключевые преимущества и недостатки каждой из четырех выделенных групп методов, а также возможности применения модифицированных методов машинного обучения для прогнозирования российской инфляции. В Табл. 1 и 2 Приложения (см. онлайн-версию статьи) приведено описание методологии и результатов российских и зарубежных исследований по прогнозированию инфляции.

В научной литературе выделяется ряд эконометрических методов прогнозирования инфляции, первым из которых является AR первого или более высокого порядка лага<sup>2</sup>. Более высокая точность прогнозирования инфляции на основе AR в сравнении с более продвинутыми моделями выявлена на статистических данных по Великобритании (Joseph et al., 2021), России (Styrin, 2019), США (Stock and Watson, 2007; Faust and Wright, 2013). Вторым видом эконометрических методов прогнозирования инфляции является модель RW и ее модификации в духе работы Atkenson and Ohanian (2001), в которой на статистических данных по США авторы получают более точные в сравнении с моделью кривой Филлипса (Phillips Curve, PC) прогнозы инфляции на последующий год через усреднение инфляции за четыре предыдущих квартала (RW-AO). Также стоит упомянуть UCSV-модели. Согласно Stock and Watson (2007), одномерная UCSV в сравнении с AR предоставляет более точные прогнозы инфляции в США. Stella and Stock (2012) предложили многомерную UCSV, которая превосходит одномерные UCSV по точности прогнозирования инфляции в США. К группе эконометрических методов прогнозирования инфляции относится и PC<sup>3</sup>. Спецификация PC может отличаться в зависимости от способов оценки предельных издержек (marginal costs) и степени недоиспользования ресурсов (slackness) в экономике: отклонения фактического значения безработицы от естественного уровня (Stock and Watson, 2008) и отклонения выпуска от потенциального уровня (Stock and Watson, 1999). Вторым критерием, по которому различаются модели PC, является характер включаемых в модель инфляционных ожиданий: 1) адаптивные, назадсмотрящие (backward-looking) ожидания; 2) рациональные, впередсмотрящие (forward-looking); 3) гибридные, включающие как адаптивные, так и рациональные ожидания. Ряд исследований показывают, что PC в сравнении с RW-AO точнее прогнозирует инфляцию в еврозоне (Banbura and Bobeica, 2023) и США (Stock and Watson, 1999). Одним из недостатков PC является линейная взаимосвязь между инфляцией и предельными издержками. Данный пробел восполняют Harding et al. (2023), которые получают более точные прогнозы инфляции в США при использовании нелинейной PC в сравнении с линейной. Следующей модификацией PC является включение в модель разрыва валютного курса вместо разрыва безработицы. Styrin and Zamulin (2012) разработали теоретическую модель PC, которая основана на отклонении валютного курса от его фундаментального значения, и эмпирически обосновали применимость предложенной PC при оценке инфляции в странах – экспортерах сырьевых товаров (Австралия, Канада, Новая Зеландия, Норвегия, Россия, ЮАР). Помимо PC отдельного внимания заслуживают альтернативные эконометрические методы, которые позволяют получить более точные прогнозы российской инфляции в сравнении с эталонными методами: метод комбинирования векторных

<sup>2</sup> В работах Styrin (2019) и Joseph et al. (2021) порядок лага в модели авторегрессии определяется байесовским информационным критерием (Bayesian Information Criterion, BIC), а в работе Stock and Watson (2007) – информационным критерием Акаике (Akaike Information Criterion, AIC).

<sup>3</sup> См. Phillips (1958), Fisher (1973).

авторегрессий и линейных регрессий (Андреев, 2016), модели с данными смешанной частоты (Mixed Data Sampling, MIDAS) (Третьяков и Фокин, 2021), модель векторной коррекции ошибок (Перевышин, 2022).

Второй группой методов прогнозирования инфляции являются *байесовские методы*. Применимость байесовских методов при прогнозировании инфляции объясняется возможностью внедрения предпосылки о том, что взаимосвязь между инфляцией и безработицей, показателями производственного сектора может меняться во времени из-за структурных сдвигов в экономике<sup>4</sup>, возникновение которых влияет на стабильность оценок РС (Stock and Watson, 2010). Geerolf (2020) отмечает, что взаимосвязь между инфляцией и безработицей в Великобритании и США сохранялась при Бреттон-Вудской системе фиксированного валютного курса, но с переходом этих стран к режиму плавающего валютного курса предпосылки исходной РС перестали соблюдаться в результате структурного сдвига. Maheu and Gordon (2008) заключают, что внедрение предпосылки о меняющемся характере параметров РС повышает точность прогнозирования инфляции в США по сравнению с традиционной РС. Полученный ими результат объясняется тем, что в модифицированной РС учитывается структурный сдвиг в форме шока цен на нефть 1970 г. Дальнейшее повышение точности байесовских методов может быть достигнуто за счет предпосылки о наличии множества структурных сдвигов в сравнении с предпосылкой о единственном структурном сдвиге, что подтверждают Koop and Potter (2007) на примере прогнозирования инфляции в США. Следующей модификацией байесовских методов прогнозирования является байесовское усреднение моделей (Bayesian Model Averaging, BMA). Согласно Kapetanios et al. (2008), BMA по сравнению с AR позволяет точнее прогнозировать инфляцию в Великобритании. Дальнейшим этапом в развитии байесовских методов прогнозирования является разработка динамического усреднения моделей (Dynamic Model Averaging, DMA) и динамического выбора моделей (Dynamic Model Selection, DMS). Koop and Korobilis (2012) показывают, что DMA и DMS в сравнении с AR и RW точнее прогнозируют инфляцию в США. На российских данных, напротив, не выявлено прогнозного преимущества DMA перед AR(1) (Styrin, 2019).

Третьей группой методов прогнозирования инфляции являются *ансамблевые методы* (ensemble methods), такие как градиентный бустинг, случайный лес (random forest), бэггинг (bagging). Многочисленные исследования показывают, что применение вышеуказанных ансамблевых методов позволяет точнее, чем AR и RW, прогнозировать инфляцию в Бразилии (Garcia et al., 2017; Araujo and Gaglianone, 2023), Великобритании (Chakraborty and Joseph, 2017), еврозоне (Zeng, 2017), России (Baybuza, 2018; Mamedli and Shibitov, 2021; Shulyak, 2022; Semiturkin and Shevelev, 2023), США (Inoue and Kilian, 2008; Medeiros et al., 2019; Coulombe et al., 2022). Kohlscheen (2022) подтверждает этот вывод на выборке из 20 развитых стран, Medeiros et al. (2023) – на обширной выборке развитых и развивающихся экономик (91 страна).

К четвертой группе методов, применяемых при прогнозировании инфляции, относятся *нейронные сети* (neural networks). Преимущество полносвязных нейронных

<sup>4</sup> Гафаров (2011) при оценке темпов роста российского индекса потребительских цен (ИПЦ) на основе РС с учетом структурных сдвигов получает, что в период с I квартала 1999 г. по II квартал 2007 г. коэффициент при уровне безработицы имеет положительный знак, а в период с III квартала 2007 г. по I квартал 2010 г. – отрицательный. Кроме того, согласно тесту Кванда – Эндрюса, отвергается нулевая гипотеза о стабильности коэффициентов при безработице.

сетей (Artificial Neural Networks, ANN) и RNN перед эталонными методами подтверждается на примере прогнозирования инфляции в Бразилии (Araujo and Gaglianone, 2023), России (Pavlov, 2020), США (Nakamura, 2005; Paranhos, 2021; Almosova and Andresen, 2023; Barkan et al., 2023). Longo and Soltanieh-ha (2023) показывают, что модель LSTM превосходит РС при прогнозировании инфляции в США. Aanes and Gullien (2018), сравнивая качество прогнозирования инфляции в Норвегии с помощью CNN и RW, приходят к выводу, что CNN позволяет получить более точный прогноз. Coulombe (2022) предлагает нейронную сеть, спецификация которой аналогична РС, но помимо разрыва выпуска и инфляционных ожиданий в ней принимаются во внимание цены на нефть и природный газ. Однако с учетом пандемического шока 2020 г. предложенная нейронная сеть оказалась не способна превзойти AR(4) по точности прогнозирования инфляции в США. Joseph et al. (2021) обнаруживают, что при прогнозировании инфляции в Великобритании точность ANN не выше точности AR(2).

Резюмируя результаты существующих исследований, отметим преимущества и недостатки каждой группы методов прогнозирования. Преимуществом *эконометрических методов* является возможность теоретически обосновать спецификации моделей, а недостатком – частотный подход (frequency-based approach) к оценке моделей, что не позволяет учесть меняющуюся во времени взаимосвязь между переменными из-за структурных сдвигов в экономике<sup>5</sup>. Указанный недостаток эконометрических методов преодолевается *байесовскими методами* прогнозирования, но только при подборе таких априорных предположений, которые соответствуют процессу генерации данных (data generating process). Вторым преимуществом байесовских методов является возможность комбинирования моделей<sup>6</sup> (например, DMA).

К преимуществам ансамблевых методов (например, модели случайного леса), во-первых, следует отнести возможность решения проблемы переобучения через агрегирование (например, усреднение) прогнозов множества моделей (деревьев решений). Во-вторых, преимущество ансамблевых методов заключается в возможности идентифицировать наиболее значимые предикторы через бутстрап-выборки, а также выявлять нелинейные взаимосвязи<sup>7</sup> между переменными через комбинацию кусочно-линейных функций.

Преимущество нейросетевых методов прогнозирования заключается в возможности точно аппроксимировать непрерывные функции. Однако нейросетевые методы также не лишены недостатков. Во-первых, ANN не способны учесть временную зависимость показателей, а именно тот факт, что статистические данные за предыдущий момент времени являются важными предикторами для прогноза на

<sup>5</sup> Отметим, что недостатком такого эконометрического метода, как MIDAS, является высокая параметризация с увеличением лагов объясняющих переменных, а преимуществом – возможность применения высокочастотных данных при прогнозировании низкочастотных показателей.

<sup>6</sup> Андреев (2016) указывает, что к преимуществам метода комбинирования прогнозов следует отнести: 1) возможность использования большого массива данных без их одновременного учета в единственной модели; 2) гибкость по отношению к структурным сдвигам за счет пересмотра весов разных прогнозов; 3) сокращение случайной ошибки посредством усреднения прогнозов. В то же время недостатком метода является неопределенность весов моделей в результирующем прогнозе, так как оценка этих весов может стать источником дополнительных ошибок (Андреев, 2016).

<sup>7</sup> Garratt et al. (2003) полагают, что нелинейная взаимосвязь между инфляцией и уровнем занятости может быть обусловлена неопределенностью в экономике.

последующий период. Этот пробел восполняют RNN, которые последовательно обрабатывают информацию: вектор скрытого представления подается на вход следующему блоку RNN. Второе преимущество RNN заключается в выявлении долгосрочных закономерностей между переменными, что осуществляется посредством ячейки памяти (memory cell). Преимущество CNN заключается в обработке многомерных матриц (тензоров) через агрегирующие фильтры, ядра (kernel) и объединяющие слои (pooling layers).

Стоит отметить, что, несмотря на преимущества CNN и RNN (а также LSTM и GRU) по сравнению с ANN, в исследованиях, посвященных применению методов машинного обучения для прогнозирования инфляции в России, вышеперечисленным модифицированным моделям не уделяется достаточного внимания. Настоящая работа восполняет данный пробел в литературе<sup>8</sup>. Более подробно обработку информации с помощью методов машинного обучения мы рассматриваем в Разделе 3, который посвящен методологии нашего исследования. Наряду с выбором инструментального подхода к прогнозированию инфляции важным фактором точности прогноза является содержательный подбор макроэкономических переменных, что обсуждается в Разделе 4.

### 3. Методология исследования

Задача прогнозирования на  $h$  временных тактов вперед заключается в нахождении отображения (mapping) между входными признаками (features)  $\mathcal{X}$  и целевым вектором  $y$  (Athey and Imbens, 2019; Masini et al., 2023):

$$y_{t+h} = \mathcal{F}_M(\mathcal{X}_t) + \epsilon_{t+h}, \quad (1)$$

где  $\mathcal{F}_M(\cdot)$  – модель  $M$  из семейства измеримых<sup>9</sup> функций  $\mathcal{F}$  (Masini et al., 2023);  $\epsilon_{t+h}$  – непрогнозируемая переменная.

Выборочная совокупность  $\mathcal{S}$  включает тренировочное  $\mathcal{L}$  и тестовое  $\mathcal{T}$  множества ( $\{\mathcal{L}, \mathcal{T}\} \in \mathcal{S}$ ). В настоящей работе оцениваются два вида отображений:  $\mathcal{F}_{ensemble}(\cdot)$  – ансамблевые модели: бэггинг, случайный лес, модифицированные версии градиентного бустинга – AdaBoost, CatBoost, LightGBM, Sklearn, XGBoost;  $\mathcal{F}_{networks}(\cdot)$  – нейронные сети: ANN, CNN, RNN, LSTM, GRU ( $\{\mathcal{F}_{ensemble}, \mathcal{F}_{networks}\} \in \mathcal{F}$ ). Каждое отображение включает матрицу признаков  $\mathcal{X}$  и вектор гиперпараметров  $\mathcal{G}$  ( $\{\mathcal{G}_{ensemble}, \mathcal{G}_{networks}\} \in \mathcal{G}$ ).

Поскольку, как уже отмечалось в Разделе 2, в исследованиях на российских данных недостаточно внимания уделяется применению CNN и RNN, а также модифицированным версиям градиентного бустинга, мы приводим ниже описание методов машинного обучения, которые ранее не использовались в российских работах по прогнозированию инфляции (Baybuza, 2018; Mamedli and Shibitov, 2021; Pavlov, 2022; Shulyak, 2022). К данным методам машинного обучения относятся: 1) бэггинг; 2) AdaBoost; 3) LightGBM; 4) CatBoost; 5) CNN; 6) RNN; 7) LSTM; 8) GRU.

<sup>8</sup> Стоит отметить, что задача прогнозирования инфляции может дополняться оценкой инфляционных ожиданий посредством текстовой информации на основе нейронных сетей Transformer, что является предметом для дальнейших исследований.

<sup>9</sup> Важность концепции измеримости функции заключается в том, что исследуемые показатели информационного множества представляют собой случайные величины, функция распределения которых определена в классе измеримых подпространств  $\mathcal{A}$  ( $\sigma$ -алгебр) пространства элементарных исходов  $\Omega$ .

1) *Алгоритм бэггинга* (Breiman, 1996): тренировочное множество  $\mathcal{L}$  разделяется на последовательность бутстрап-подмножеств с возвращением (bootstrap replications with replacement)  $\{\mathcal{L}^{(B)}\}$ . На основе отдельного подмножества оценивается отображение между предикторами  $x$  и целевым вектором  $y$  на основе базовой регрессии  $\{\varphi(x, \mathcal{L}^{(B)})\}$ . Итоговый результат бэггинга основан на усреднении оценок по всем бутстрап-подмножествам:

$$\varphi_B(x) = \mathbb{E}_B[\varphi(x, \mathcal{L}^{(B)})], \quad (2)$$

где  $\varphi_B$  – агрегированная функция;  $\mathbb{E}_B$  – оператор математического ожидания по бутстрап-выборкам.

В настоящей работе базовой моделью (base regressor) является дерево решений (decision tree), что по умолчанию задано в библиотеке *sklearn*<sup>10</sup>.

2) *Адаптивный бустинг AdaBoost* (Freund and Schapire, 1997): при информационном множестве  $\{(x_i, y_i)\}_{i=1}^N$  задаются начальные веса  $w_i^t$  для обновления оценок базовой модели обучения (weak learner) на временном шаге  $t$ ,  $h_t(x_i) : X \rightarrow [0, 1]$ <sup>11</sup>. На основе весов  $w^t$  определяется вектор нормированных распределений  $p^t = w^t / \sum_{i=1}^N w_i^t$ , а также регрессионные ошибки моделей:

$$\epsilon_t = \sum_{i=1}^N p_i^t |h_t(x_i) - y_i|.$$

При этом ошибки моделей варьируются в интервале  $\epsilon_t \in [0, 1]$ . Механизм адаптации заключается в обновлении поправочных коэффициентов  $\beta_t = \epsilon_t / (1 - \epsilon_t)$ , с учетом которых обновляются оценки моделей:

$$w_i^{t+1} = w_i^t \beta_t^{1 - |h_t(x_i) - y_i|}.$$

Итоговый результат представляет собой

$$h_f(x) = \begin{cases} 1, & \text{если } \sum_{t=1}^T \left( \log \left( \frac{1}{\beta_t} \right) h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \log \left( \frac{1}{\beta_t} \right) \right), \\ 0 & \text{иначе.} \end{cases} \quad (3)$$

Аналогично алгоритму бэггинга в качестве базовой конструкции выступает модель дерева решений, что по умолчанию задано в библиотеке *sklearn*<sup>12</sup>.

3) *Градиентный бустинг LightGBM* (Ke et al., 2017): преимущество алгоритма заключается в увеличении глубины дерева на уровне листьев (leaf-wise tree growth), а не на уровне переменных (level-wise tree growth)<sup>13</sup>. В базовой модели дерева решений значимость  $j$ -го предиктора оценивается приростом дисперсии (variance gain):

$$V_{j|O}(d) = \frac{1}{n_O} \left( \frac{\left( \sum_{\{x_i \in O: x_{ij} \leq d\}} g_i \right)^2}{n_{l|O}^j} + \frac{\left( \sum_{\{x_i \in O: x_{ij} > d\}} g_i \right)^2}{n_{r|O}^j} \right), \quad (4)$$

<sup>10</sup> См. <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.BaggingRegressor.html>

<sup>11</sup> Отметим, что параметр  $h_t$  в алгоритме градиентного бустинга участвует в итеративном обновлении оценок моделей  $F^t = F^{t-1} + \alpha h_t$ , где  $\alpha$  – темп обучения (learning rate). Подробнее про градиентный бустинг см. в работе Friedman (2001).

<sup>12</sup> См. <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.AdaBoostRegressor.html>

<sup>13</sup> См. <https://lightgbm.readthedocs.io/en/latest/Features.html>

где  $d$  – пороговое значение наблюдения, при котором выборка разделяется на два класса: правый и левый узлы (left and right nodes);  $g_i$  – антиградиент функции потерь по параметрам  $i$ -й модели;  $n_o = \sum I[x_i \in O]$  – количество наблюдений тренировочного множества;  $n_{l|O}^j(d) = \sum I[x_i \in O : x_{ij} \leq d]$  – наблюдения первого класса, в котором величина  $j$ -го признака не больше порогового уровня  $d$ ;  $n_{r|O}^j(d) = \sum I[x_i \in O : x_{ij} > d]$  – наблюдения второго класса, в котором величина  $j$ -го признака превышает пороговый уровень  $d$ .

Модификация LightGBM заключается в фильтрации наблюдений с малым вкладом в обновление весов моделей, то есть наблюдений с малыми величинами градиентов<sup>14</sup> (small gradients). Наблюдения сортируются в убывающем порядке по абсолютному значению градиента, из которых выбираются  $a \times 100\%$  (множество  $A$ ), а из остальной части выборки с наименьшим значением градиента (множество  $A^c$ ) отбираются  $b \times 100\%$  наблюдений (множество  $B$ ), градиент которых умножается на поправочный коэффициент  $(1 - a)/b$ . Модифицированный прирост дисперсии имеет вид

$$\tilde{V}_j(d) = \frac{1}{n} \left( \frac{\left( \sum_{\{x_i \in A_l\}} g_i + \frac{1-a}{b} \sum_{\{x_i \in B_l\}} g_i \right)^2}{n_l^j(d)} + \frac{\left( \sum_{\{x_i \in A_r\}} g_i + \frac{1-a}{b} \sum_{\{x_i \in B_r\}} g_i \right)^2}{n_r^j(d)} \right), \quad (5)$$

где  $B_r = \{x_i \in B : x_{ij} > d\}$  и  $A_r = \{x_i \in A : x_{ij} > d\}$  – множества правого узла;  $B_l = \{x_i \in B : x_{ij} \leq d\}$  и  $A_l = \{x_i \in A : x_{ij} \leq d\}$  – множества левого узла.

Тем самым, согласно алгоритму LightGBM, оценка прироста дисперсии и последующее увеличение глубины дерева решений осуществляются на выборке с меньшим количеством наблюдений, а не на всем тренировочном множестве. Техническая реализация алгоритма LightGBM в настоящей работе основана на библиотеке *lightgbm*<sup>15</sup>.

4) *Градиентный бустинг CatBoost* (Categorical Boosting; см. Dorogush et al., 2018; Prokhorenkova et al., 2018). Модификация алгоритма CatBoost заключается в том, что на каждой  $t$ -й итерации обучение дерева решений  $T_t$  (из общего количества деревьев  $I$ ) происходит на случайно перемешанной (random permutation) выборке  $\sigma_t$  из множества комбинаций  $\{\sigma_1, \sigma_2, \dots, \sigma_s\}$ . Прогнозы вспомогательной модели (supporting model)  $M_{r,j}(i)$  для следующего  $i$ -го наблюдения составляются на основе предыдущих  $j$ -х величин выборки  $\sigma_t$ . Ошибки моделей  $r^t(x_k, y_k)$  вычисляются посредством модели предыдущей итерации.

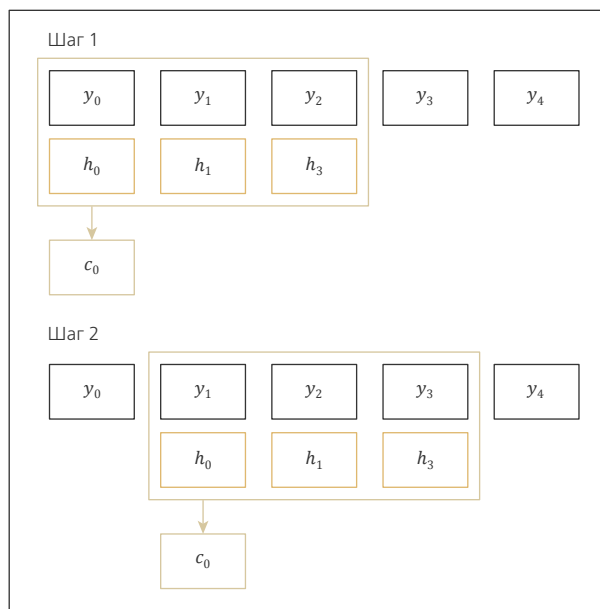
5) *CNN*. В нашем исследовании мы применяем одномерные CNN (1D-CNN)<sup>16</sup>, свертка которых осуществляется в одном измерении, что проиллюстрировано на Рис. 1.

<sup>14</sup> Градиент – это вектор частных производных функции потерь (loss function) по отдельным параметрам.

<sup>15</sup> См. <https://lightgbm.readthedocs.io/en/latest/Python-Intro.html>

<sup>16</sup> Применение двумерных нейронных сетей для прогнозирования российской инфляции может быть предметом дальнейших исследований.



**Рисунок 1. Механизм работы одномерных CNN**

Источник: составлено автором

Последовательность случайных величин  $(y_0, y_1, y_2, y_3, y_4)$  отражает прогнозируемый показатель. Вектор свертки характеризуется последовательностью  $(h_0, h_1, h_2)$ . Первым результатом свертки является вектор  $c_0$ , который представляет собой взвешенную сумму произведений:  $c_0 = y_0 h_0 + y_1 h_1 + y_2 h_2$ . На втором шаге свертки мы получаем следующий результат:  $c_1 = y_1 h_0 + y_2 h_1 + y_3 h_2$ . Одним из примеров 1D-CNN является нейронная сеть WaveNet, предназначенная для генерации аудио (van den Oord et al., 2016).

Для задачи прогнозирования инфляции в настоящей работе входной тензор (многомерная матрица) состоит из трех элементов: 1) количество подвыборок (batches); 2) число временных шагов/лагов (time steps); 3) количество предикторов (features). Из каждой подвыборки входного тензора извлекаются матрицы меньшей размерности на основе скользящих фильтров (sliding filters)<sup>17</sup>. Отфильтрованные трехмерные матрицы преобразовываются в одномерные векторы посредством тензорного произведения (Кroneкера)  $\otimes$  со сверточными ядрами (convolutional kernels). Вычисленное двумерное произведение вновь преобразовывается в трехмерный формат.

6) RNN. На каждом цикле обучения RNN принимает входные параметры, а также формирует память посредством фиксации весовой матрицы узлов (state) между подвыборками, что позволяет сформировать результирующий вектор (output) на следующей итерации.

<sup>17</sup> Зачастую матричная размерность скользящих фильтров составляет  $3 \times 3$ ,  $5 \times 5$ .

Дальнейшее описание RNN, LSTM и GRU соответствует Chung et al. (2014). Модель RNN фиксирует весовую матрицу узлов скрытых слоев (hidden state) в каждый момент времени на основе результата обучения предыдущего цикла. Так, для последовательности чисел  $\{x_t\}_{t=1}^T$  обновление скрытого состояния происходит следующим образом:

$$h_t = \begin{cases} 0, & \text{если } t = 0, \\ \phi(h_{t-1}, x_t) & \text{иначе,} \end{cases} \quad (6)$$

где  $\phi(\cdot)$  – нелинейная функция;  $h_t = g(Wx_t + Uh_{t-1})$  – частный пример функции обновления состояния;  $W, U$  – матрицы весов.

7) *LSTM*. Эти модели были разработаны для анализа последовательных объектов (например, прогнозирования смыслового завершения предложения на основе предыдущих слов). LSTM является модифицированной версией RNN. Модификация заключается во внедрении пропускных механизмов (gated mechanisms) потока информации<sup>18</sup>. На каждом цикле  $t$  обучения LSTM формирует результат  $j$ -го нейрона  $h_t^j$  следующим образом:

$$\begin{cases} h_t^j = o_t^j \times \tanh(c_t^j), \\ o_t^j = \sigma(W_o x_t + U_o h_{t-1} + V_o c_t^j), \\ c_t^j = f_t^j c_{t-1}^j + i_t^j \tilde{c}_t^j, \\ \tilde{c}_t^j = \tanh(W_c x_t + U_c h_{t-1}), \\ f_t^j = \sigma(W_f x_t + U_f h_{t-1} + V_f c_{t-1}^j), \\ i_t^j = \sigma(W_i x_t + U_i h_{t-1} + V_i c_{t-1}^j), \end{cases} \quad (7)$$

где  $j$  – индекс узла нейронной сети;  $t$  – индекс цикла обучения;  $o_t^j$  – результирующий пропускной механизм (output gate);  $c_t^j$  – ячейка памяти (memory cell);  $\tilde{c}_t^j$  – обновление ячейки памяти;  $f_t^j$  – пропускной механизм отсеивания информации (forget gate);  $i_t^j$  – пропускной механизм фиксации информации (input gate);  $\tanh$  – гиперболический тангенс;  $\sigma$  – логистическая сигмоидальная функция (logistic sigmoid function).

8) *GRU*. Прогноз в модели GRU формируется следующим образом:

$$\begin{cases} h_t^j = (1 - z_t^j) h_t^j + z_t^j \tilde{h}_t^j, \\ \tilde{h}_t^j = \tanh(W x_t + U(r_t \odot h_{t-1})), \\ z_t^j = \sigma(W_z x_t + U_z h_{t-1}), \\ r_t^j = \sigma(W_r x_t + U_r h_{t-1}), \end{cases} \quad (8)$$

где  $j$  – индекс узла нейронной сети;  $t$  – индекс цикла обучения;  $z_t^j$  – пропускной механизм обновления информации (update gate);  $\odot$  – поэлементное умножение матриц;  $\tilde{h}_t^j$  – потенциальный результат (candidate activation);  $r_t^j$  – пропускной механизм обновления информации (reset gate). Тем самым ячейка памяти  $c_t^j$  LSTM заменяется пропускным механизмом обновления информации  $r_t^j$  в GRU.

<sup>18</sup> Подробное описание RNN и LSTM см. в работах Hochreiter and Schmidhuber (1997) и Karpathy et al. (2015).

Далее мы описываем сетку гиперпараметров и способы выявления оптимальной архитектуры методов машинного обучения. Более полная информация по априорным гиперпараметрам ансамблевых моделей и нейронных сетей приведена в Табл. 3–4 Приложения, здесь мы перечисляем только ключевые характеристики моделей.

Начальная сетка гиперпараметров для ансамблевых методов и нейронных сетей задается на основе результатов научных работ: модели случайного леса (Baybuza, 2018); модели градиентного бустинга (Shulyak, 2022); ANN (Pavlov, 2020); CNN (Aanes and Gullien, 2018); GRU (Barkan et al., 2023); LSTM (Paranhos, 2021). Нейронные сети обучаются на основе стохастической оптимизации (Kingma and Ba, 2017).

Оптимальные гиперпараметры моделей выбираются кросс-валидацией при помощи алгоритма разбиения временного ряда TimeSeriesSplit из библиотеки *sklearn*<sup>19</sup>. Подробнее опишем данный алгоритм. На первом шаге модели обучаются и тестируются на 20% выборочной совокупности, а на последующих шагах – на тренировочном множестве, объем которого итеративно увеличивается на 20%. Оптимальные гиперпараметры выбираются на том количестве разбиений, при котором достигается минимальная среднеквадратичная ошибка (Mean Squared Error, MSE). Оптимальные гиперпараметры ансамблевых и нейросетевых моделей приведены в Табл. 7, 9, 11 и 13 Приложения.

Результаты моделей интерпретируются с использованием вектора Шепли (см. Lundberg and Lee, 2017). Из всего множества факторов  $F$  выбираются подмножества признаков  $S \subseteq F$ , на основе которых обучаются два вида моделей  $f$ : с учетом  $i$ -го предиктора  $f_{S \cup \{i\}}$ ; без учета  $i$ -го предиктора  $f_S$ . Итоговое значение Шепли (Shapley value) вычисляется посредством среднего значения разности результатов вышеприведенных двух видов моделей, взвешенного по всем комбинациям подмножеств признаков:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|! (|F| - |S| - 1)!}{|F|!} [f_{S \cup \{i\}}(x_{S \cup \{i\}}) - f_S(x_S)], \quad (9)$$

где  $\phi_i$  – вклад предиктора  $i$  в изменение итогового прогноза;  $x_S$  – значения предикторов подмножества  $S$ ;  $|F|$  – количество ненулевых элементов множества  $F$ ;  $S \subseteq F \setminus \{i\}$  – подмножество признаков без учета предиктора  $i$ .

Для технической реализации алгоритма используется библиотека *shap*<sup>20</sup>.

#### 4. Описание статистических данных

Наряду с методологическим аспектом прогнозирования инфляции важным является подбор макроэкономических переменных, которые позволяют точнее прогнозировать ценовую динамику. Мы составили пять групп предикторов инфляции на основе двух критериев: 1) наличие теоретических механизмов взаимосвязи между инфляцией и показателями отдельных групп; 2) повышение точности прогнозирования инфляции за счет включения показателей отдельных групп. Более подробная информация по используемым показателям приведена в Табл. 5 Приложения.

<sup>19</sup> См. [https://scikit-learn.org/stable/modules/generated/sklearn.model\\_selection.TimeSeriesSplit.html](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.TimeSeriesSplit.html)

<sup>20</sup> См. <https://shap.readthedocs.io/en/latest/index.html>

К *первой группе* относятся показатели рынка труда: уровень безработицы (Phillips, 1958; Fisher, 1973) и номинальная заработная плата (Styrin, 2019). Согласно РС, снижение уровня безработицы ассоциируется с ростом номинальной заработной платы. Повышение номинальной заработной платы приводит к ускорению инфляции. При росте цен возникает необходимость повышения номинальной заработной платы, то есть раскручивается спираль «заработная плата – цены». Также отклонение фактического уровня безработицы от ее естественного уровня связано с циклическими колебаниями выпуска, которые создают проинфляционный эффект (Орлов и Постников, 2020).

*Вторая группа* состоит из показателей агрегированного спроса: объема платных услуг населению и оборота розничной торговли (Перевишин, 2022). В обобщенной РС предполагается, что отклонение инфляции от ее ожидаемого значения (так называемая неожиданная инфляция) связано с отклонением выпуска от его потенциального уровня (разрывом выпуска). Формализовать данную взаимосвязь позволяет закон Оукена, согласно которому разрыв выпуска определяется отклонением фактической безработицы от ее естественного уровня (циклической безработицей).

В *третью группу* мы включили показатели внутреннего финансового сектора: кредиты, выданные частному сектору (Styrin, 2019); денежные агрегаты (Forni et al., 2003; Pavlov, 2020); доходность государственных облигаций (Stock and Watson, 2003; Estrella, 2005). Одним из объяснений возможности применения показателей финансового сектора в прогнозировании инфляции является гипотеза ожиданий, согласно которой процентные ставки отражают ожидания экономических агентов относительно денежно-кредитной политики, оказывающей непосредственное влияние на инфляцию и выпуск<sup>21</sup>.

В *четвертой группе* учитываются показатели, характеризующие условия денежно-кредитной и бюджетно-налоговой политики: процентная ставка MIACR (Третьяков и Фокин, 2021); процентная ставка RUONIA (Перевишин, 2022); объем государственного долга (Cochrane, 2001; Sims, 2010; Araujo and Gaglionne, 2023); государственные расходы (Kinlaw et al., 2023), а также доходы государственного бюджета. Трансмиссионный механизм заключается в том, что стимулирующая денежно-кредитная политика и бюджетно-налоговая политика создают проинфляционный эффект через рост деловой активности в результате смягчения условий на финансовом рынке и увеличения агрегированного спроса.

К *пятой группе* относятся инфляционные ожидания (Coibion et al., 2018; Angelico et al., 2022; Shulyak, 2022; Kinlaw et al., 2023). При установлении номинальных заработных плат экономические агенты принимают во внимание ожидаемую инфляцию, так как они заинтересованы в реальной покупательной способности получаемых доходов. РС, в которой учитываются инфляционные ожидания, впервые была предложена в работах Friedman (1968) и Phelps (1968)<sup>22</sup>.

<sup>21</sup> См., например, работу Estrella (2005), в которой автор объединяет правило Тейлора, уравнение Фишера (для номинальной ставки процента) и РС, для того чтобы аналитически вывести взаимосвязь между будущей инфляцией (будущим выпуском) и дифференциалом долгосрочных и краткосрочных процентных ставок.

<sup>22</sup> Согласно расширенной РС, существует взаимосвязь между отклонением фактической инфляции  $\pi_t$  от ее ожидаемого значения  $\pi_t^e$  (то есть  $\pi_t - \pi_t^e$ ) и отклонением фактической безработицы  $u_t$  от ее естественного уровня  $u^n$  (то есть  $u_t - u^n$ ).

В *шестую группу* мы включили глобальные факторы: валютные курсы (Андреев, 2019; Хабибуллин, 2019; Saul, 2021; Semiturkin and Shevelev, 2023); цены на энергоносители и продовольственные товары (Kiselev and Zhivaykina, 2020; Абрамов и др., 2022); инфляцию в еврозоне (Nagy and Tengey, 2018), США (Ditzen and Ravazzolo, 2022). Прямым каналом влияния глобальной экономики на инфляцию внутри малой открытой экономики является рост цен импортных товаров, а косвенным каналом – рост мирового спроса на товары, производимые малой открытой экономикой, что создает инфляционное давление (Абрамов и др., 2022). Также рост цен на энергоносители приводит к ускорению инфляции в странах – импортерах сырья через рост издержек производства, а в странах – экспортерах сырья – через повышение совокупного спроса в результате роста экспортной выручки.

Целевой переменной прогнозирования является ИПЦ в % к соответствующему месяцу предыдущего года (г/г). Все используемые факторы берутся с одним лагом, приведены к формату г/г, а также стандартизированы согласно формуле<sup>23</sup> (11):

$$z_t = \frac{x_t - x_{\min}}{x_{\max} - x_{\min}}, \quad (11)$$

где  $x_t$  – исходное значение переменной в период  $t$ ;  $x_{\min}$  – минимальное значение переменной в течение тренировочного множества;  $x_{\max}$  – максимальное значение переменной в течение тренировочного множества.

Источниками статистических данных являются: 1) периодическое издание «Краткосрочные экономические показатели», публикуемое Федеральной службой государственной статистики (Росстат); 2) информация об исполнении федерального бюджета, публикуемая Министерством финансов России; 3) динамика процентных ставок, публикуемая Банком России; 4) цены на сырьевые товары из базы данных Федерального резервного банка Сент-Луиса FRED. С учетом доступности статистических данных<sup>24</sup> исследуемый период составляет январь 2012 г. – апрель 2023 г. Тренировочное множество охватывает период с января 2012 г. по июль 2019 г., что составляет 67% выборочной совокупности. На тестовое множество приходится 33% наблюдений выборки – с августа 2019 г. по апрель 2023 г.

После группировки факторов и формирования тренировочного и тестового множеств следующим этапом предобработки данных является тензорное представление выборочной совокупности согласно следующему алгоритму.

1. Выбираем тип информационного множества: 1) одномерный случай, состоящий только из лаговых величин инфляции,  $\mathcal{L}^{cpi}$ ; 2) многомерный случай, включающий дополнительные макроэкономические предикторы инфляции,  $\mathcal{L}^{all}$ .
2. Задаем порядок временных лагов предикторов  $\ell \in \{12, 24, 36, 48\}$ <sup>25</sup> и горизонт прогнозирования  $\mathcal{h} \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ .

<sup>23</sup> См. <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html>

<sup>24</sup> В частности, информация о показателях бюджетно-налоговой политики доступна с января 2011 г., что с учетом приведения факторов в формат г/г определяет начало рассматриваемого периода (январь 2012 г.).

<sup>25</sup> Дальнейшие результаты мы приводим при  $\ell = 24$ , так как это значение обеспечивает наиболее точный прогноз эталонной модели AR(1); при  $\ell \in \{12, 36, 48\}$  ошибка прогноза AR(1) выше. Таким образом, мы сравниваем точность прогнозирования методами машинного обучения с точностью лучшего варианта AR(1).

3. Двумерный массив (месяц  $\times$  предиктор) приводится к трехмерной матрице: количество последовательностей (batches), порядок временных лагов (time series step), количество предикторов<sup>26</sup>. На Рис. 2 мы привели иллюстративный пример. В частности, при выборе  $\mathcal{L}^{cpi}$ ,  $\ell = 3$ ,  $\mathcal{h} = 2$  из выборки (6 месяцев  $\times$  1 предиктор) получается тензор из двух последовательностей, в каждой из которых три месяца приходятся на предиктор и два месяца – на целевой вектор (Рис. 2а). В то же время при выборе  $\mathcal{L}^{all}$ ,  $\ell = 3$ ,  $\mathcal{h} = 2$  формируется тензор из двух последовательностей, каждая из которых включает двумерную матрицу из двух предикторов и трех месяцев, а также целевой вектор из двух целевых показателей и двух месяцев (Рис. 2б).

С учетом итоговой предобработки данных мы прогнозируем инфляцию в России согласно следующему алгоритму.

1. Выбираем тренировочное множество  $\mathcal{L}_{t-1}^{cpi}$ ,  $\mathcal{L}_{t-1}^{all}$  и задаем сетку гиперпараметров  $\mathcal{G}_{ensemble}$ ,  $\mathcal{G}_{networks}$ .
2. Из сетки гиперпараметров выявляем оптимальную топологию моделей через 5, 10, 15<sup>27</sup> разбиений кросс-валидации временных рядов (time series split) для получения максимально приближенного отображения (mapping) входных параметров ансамблевых моделей  $\mathcal{F}_{ensemble}$  к фактическим величинам целевого вектора (target)<sup>28</sup>. Для выявления оптимальной топологии нейронных сетей  $\mathcal{F}_{networks}$  составляется комбинация количества узлов и функций активации, при которой достигается минимальное значение MSE прогноза.
3. Используем рекурсивный подход к прогнозированию: первый прогноз заменяем на последнюю последовательность тренировочного множества, последующий прогноз основан на обновленной выборке. При одномерном информационном множестве обновляется временной ряд по инфляции, а при многомерном – по всем макроэкономическим переменным. В последнем случае решается задача прогнозирования «входная многомерная последовательность – выходная многомерная последовательность» (many-to-many sequence). Количество последовательных циклов рекурсивного прогнозирования определяется числом месяцев тестового множества.
4. Сравниваем прогнозы отдельных методов с эталонной моделью авторегрессии первого порядка (Stock and Watson, 2007; Baybuza, 2018; Kohlscheen, 2022; Shulyak, 2022) на тестовом множестве  $\mathcal{T}$  по горизонтам  $\mathcal{h} \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$  с точки зрения средней абсолютной (Mean Absolute Error, MAE) и среднеквадратической (Root Mean Square Error, RMSE) ошибок прогноза.
5. Для моделей с наиболее точными прогнозами дополнительно вычисляем ожидаемые значения вектора Шепли для содержательной интерпретации результатов.

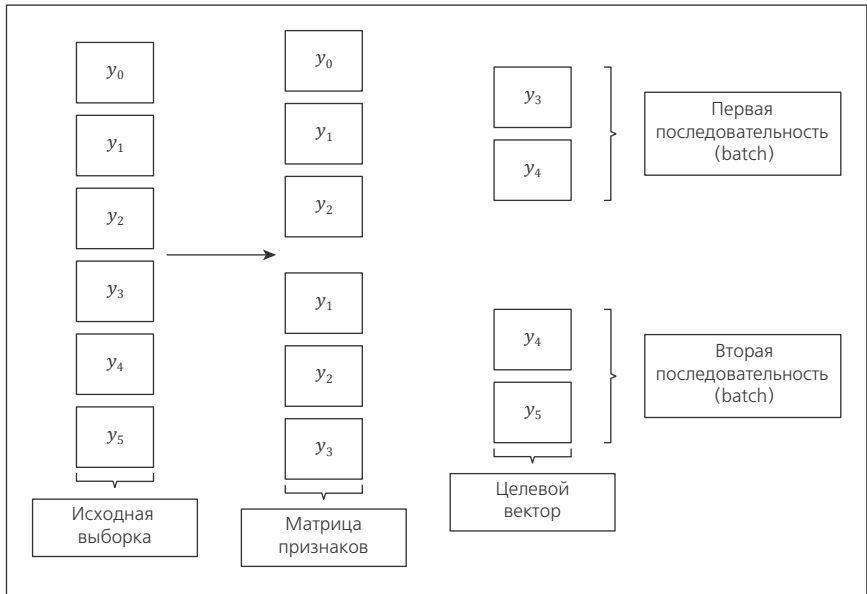
<sup>26</sup> Для формирования трехмерного тензора можно воспользоваться: 1) функцией Time Series Data From Array, <https://keras.io/api/preprocessing/timeseries/#timeseriesdatasetfromarray-function>; 2) функцией Time Series Generator, [https://www.tensorflow.org/api\\_docs/python/tf/keras/preprocessing/sequence/TimeseriesGenerator](https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/sequence/TimeseriesGenerator) (Brownlee, 2018, pp. 44–45; Chollet, 2021, pp. 285–286).

<sup>27</sup> Наиболее точные прогнозы получены при пяти разбиениях временных рядов.

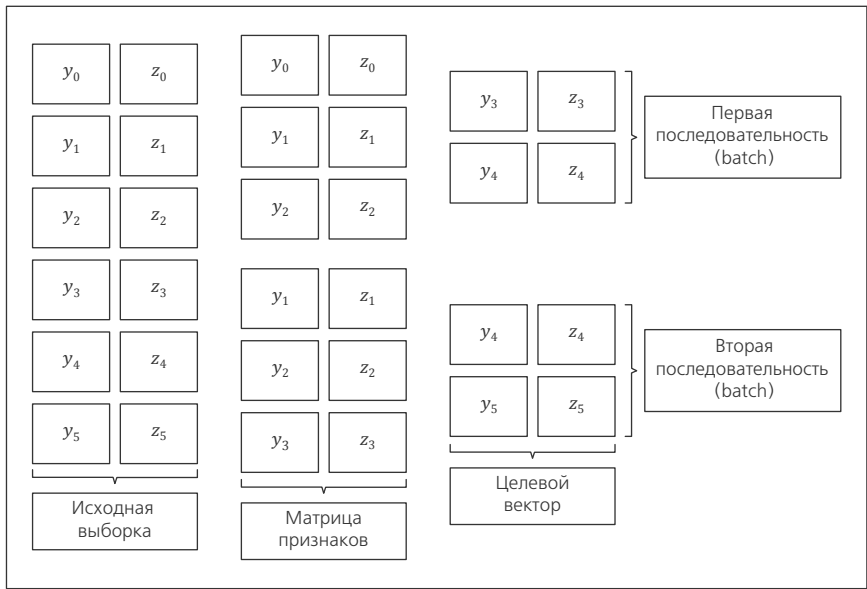
<sup>28</sup> См. [https://scikit-learn.org/stable/modules/generated/sklearn.model\\_selection.GridSearchCV.html](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html)

**Рисунок 2. Тензорное представление выборки**

**а) преобразование одномерной выборки в тензор**



**б) преобразование многомерной выборки из двух предикторов в тензор**



Примечание: горизонт прогнозирования – два месяца, длина лага – три месяца.

Источник: составлено автором

## 5. Результаты прогнозирования

В этом разделе мы рассматриваем результаты рекурсивного прогнозирования. В Табл. 6 и 8 приведены RMSE и MAE одномерных методов машинного обучения, учитывающих только лаговые величины инфляции, по отношению к RMSE и MAE эталонной модели AR(1). В Табл. 7 и 9 указаны оптимальные с точки зрения RMSE и MAE гиперпараметры одномерных методов машинного обучения. Как следует из представленных данных, все виды нейронных сетей (ANN, CNN, RNN) предоставляют более точные оценки будущей инфляции по сравнению с AR(1) на всех горизонтах прогнозирования от 1 до 12 месяцев. При этом на горизонте в 1 месяц наибольшее по сравнению с метриками точности AR(1) сокращение RMSE получено для модели нейронной сети LSTM, а MAE – для ANN. Прогнозное преимущество LSTM может объясняться пропускным механизмом последовательной обработки информации. На горизонте в 3 месяца ANN также предоставляет наиболее точные прогнозы с точки зрения RMSE и MAE. При оценке будущей инфляции на шесть месяцев вперед прогнозное преимущество оказывается у CNN. На горизонте в 12 месяцев наименьшая ошибка прогноза (как RMSE, так и MAE) получена при использовании LSTM, что может объясняться ее способностью выявлять долгосрочные закономерности. Среди одномерных ансамблевых моделей наиболее точные прогнозы в горизонтах 1 и 3 месяца получены в модели градиентного бустинга Sklearn. Однако модели случайного леса, а также градиентного бустинга AdaBoost и CatBoost не превосходят точность AR(1). Последний результат согласуется с выводами Faust and Wright (2013), Styrin (2019) и Joseph et al. (2021) о том, что в задаче прогнозирования инфляции превзойти точность эталонных моделей более сложными методами – трудно выполнимая задача.

Значения метрик точности многомерных методов машинного обучения, оцененных с учетом макроэкономических предикторов инфляции, приведены в Приложении (см. Табл. 10 и 12). Стоит отметить, что на горизонте в 1 месяц наименьшая ошибка прогноза получена при использовании модели ANN: сокращение RMSE по сравнению с бенчмарком составляет 28,23%, а MAE – 24,67%. ANN также оказывается лучшей при прогнозировании инфляции на 3 и 6 месяцев вперед. При прогнозировании инфляции на 12 месяцев вперед наиболее точным многомерным методом является LSTM. При этом наблюдается большее сокращение RMSE прогноза многомерной LSTM по сравнению с прогнозом одномерной спецификации. Таким образом, учет дополнительных макроэкономических факторов ценовой стабильности повышает точность многомерных моделей по сравнению с одномерными моделями. Среди ансамблевых моделей многомерная спецификация градиентного бустинга Sklearn характеризуется наименьшей RMSE при прогнозе инфляции на 6 и 12 месяцев вперед.

Прогнозные значения инфляции, полученные на тренировочном множестве, представлены на Рис. 6–15 в Приложении. Отметим, что как ансамблевые, так и нейросетевые модели способны идентифицировать ускорение инфляции, датированное I кварталом 2015 г. Далее на Рис. 16–25 в Приложении проиллюстрированы прогнозные значения инфляции, полученные на тестовом множестве. Преимущество ANN при прогнозе инфляции на горизонте в 1 и 6 месяцев (Рис. 17 и 21 в Приложении) объясняется тем, что данная модель среди других видов



нейронных сетей наиболее точно идентифицирует ускорение инфляции, произошедшее в феврале 2022 г.

Перейдем к интерпретации результатов наиболее точных методов машинного обучения (градиентного бустинга Sklearn и ANN<sup>29</sup>) с помощью вектора Шепли (см. Рис. 3–4 в Приложении). Согласно Sklearn, наибольший вклад в прогноз инфляции имеет ее первый лаг. Данный результат отражает адаптивный характер инфляционных ожиданий в России (Банк России, 2019). Далее следуют индекс продовольственных цен Продовольственной и сельскохозяйственной организации ООН (Food and Agriculture Organization, FAO), цена на нефть марки Brent, цена на природный газ, оборот розничной торговли продовольственными товарами и показатель инфляции в еврозоне. Это говорит о высокой степени зависимости инфляции в России от глобальных факторов, а также от внутренней потребительской активности. С учетом того что тестовая выборка охватывает ковидный период, значительный вклад потребительского спроса, измеряемого как оборот розничной торговли продовольственными товарами, объясняется превышением спроса над возможностями наращивания предложения.

Среди показателей, характеризующих денежно-кредитные условия, наибольшей информативностью для прогноза российской инфляции обладает процентная ставка RUONIA. Это ожидаемо, поскольку она отражает стоимость необеспеченных однодневных межбанковских кредитов, которая является одним из важных звеньев трансмиссионного механизма денежно-кредитной политики. Результаты реализации нашей эмпирической стратегии показывают, что курс евро к рублю входит в число факторов, вносящих существенный вклад в российскую инфляцию, что соответствует присущей российской экономике значимости эффекта переноса валютного курса в цены (Пономарев и др., 2014; Синяков и др., 2019). Данный результат объясняется тем, что обменный курс рубля включается в модель в рамках блока глобальных факторов, в том числе мировых цен на энергоносители, которые существенным образом коррелируют с переменной обменного курса рубля. В особенности высокая степень корреляции между этими переменными характерна для тренировочного множества, которое охватывает значительный период до введения в феврале 2017 г. нового бюджетного правила, ослабившего связь между ценой на нефть и обменным курсом российского рубля.

Согласно ANN, информативными переменными при прогнозе российской инфляции являются темп прироста численности безработных, что согласуется с предпосылками РС. Также в число наиболее информативных переменных входят цены на природный газ и инфляция в США и еврозоне, что может отражать импорт мировой инфляции в российскую экономику.

Для интерпретации результатов модели градиентного бустинга Sklearn мы помимо вектора Шепли используем атрибут «важность переменной»<sup>30</sup> (feature importance, см. Рис. 5 в Приложении). Наиболее важными предикторами инфляции являются индекс продовольственных цен FAO, лаг инфляции в России, инфляция

<sup>29</sup> Для Sklearn применяется метод TreeExplainer, а для ANN – DeepExplainer, см. <https://shap-lrjball.readthedocs.io/en/latest/generated/shap.TreeExplainer.html>, <https://shap-lrjball.readthedocs.io/en/latest/generated/shap.DeepExplainer.html>

<sup>30</sup> См. работы Zeng (2017), Kohlscheen (2022); пример технической реализации доступен на сайте [https://scikit-learn.org/stable/auto\\_examples/ensemble/plot\\_forest\\_importances.html](https://scikit-learn.org/stable/auto_examples/ensemble/plot_forest_importances.html)

в еврозоне, цены на нефть и природный газ, что аналогично результатам, полученным по вектору Шепли. Среди показателей налогово-бюджетной политики наибольшей информативностью обладают нефтегазовые доходы федерального бюджета, что объясняется значимой ролью выручки от экспорта энергоносителей в динамике агрегированного спроса и инфляции в России.

## 6. Выводы

В настоящей работе впервые предпринята попытка прогнозирования инфляции в России на основе модифицированных методов машинного обучения: адаптивного и упорядоченного градиентного бустинга; сверточных и рекуррентных нейронных сетей; модели долгой краткосрочной памяти. При использовании информационного множества из лаговых величин инфляции получено, что по сравнению с эталонной моделью авторегрессии первого порядка все нейросетевые модели (полносвязные, сверточные, рекуррентные нейронные сети, модель долгой краткосрочной памяти) предоставляют более точные прогнозы инфляции в горизонтах от 1 до 12 месяцев. Однако среди ансамблевых моделей прогнозное преимущество перед моделью авторегрессии первого порядка выявлено только у модели градиентного бустинга Sklearn. Данный результат согласуется с выводами Faust and Wright (2013), Styrin (2019) и Joseph et al. (2021) о том, что ряду более сложных моделей не удается превзойти бенчмарки по точности прогнозирования инфляции. При добавлении макроэкономических факторов точнее других моделей прогнозируют инфляцию полносвязные и сверточные нейронные сети. На горизонте в 12 месяцев прогнозное преимущество – у моделей долгой краткосрочной памяти, что может объясняться пропускным механизмом последовательной обработки информации и возможностью выявления долгосрочных взаимосвязей между переменными.

При интерпретации результатов наиболее точных методов машинного обучения (полносвязные нейронные сети, модель градиентного бустинга Sklearn) с помощью вектора Шепли получено, что наиболее информативными переменными при прогнозе российской инфляции являются ее лаговая величина, цена на нефть и природный газ, инфляция в еврозоне и США, нефтегазовые доходы федерального бюджета, оборот розничной торговли как продовольственных, так и непродовольственных товаров. Данные результаты свидетельствуют о влиянии на российскую инфляцию глобальных факторов, состояния федерального бюджета России, динамики внутренней потребительской активности.

Помимо моделей прогнозирования инфляции, изученных и апробированных в настоящей работе, возможна и практическая реализация ряда других методов. Как отмечают Karpetanios et al. (2008), альтернативным подходом к ВМА, который также применяется в работе Styrin (2019), является информационно-теоретическое усреднение моделей (Information-Theoretic Model Averaging, ITMA). На примере прогнозирования инфляции в Великобритании использование ITMA позволяет получить более точный прогноз по сравнению с ВМА. С учетом этого вывода Karpetanios et al. (2008) подход ITMA может быть дополнительным методом прогнозирования инфляции в России. Наряду с прогнозами по отдельным нейронным сетям возможно применение их комбинации. Так, Coulombe (2022) на примере инфляции США оценил полусферическую модель нейронной сети, которая демонстрирует

более точные прогнозы в сравнении с полносвязной нейронной сетью и РС. Наконец, подходы модифицированных алгоритмов градиентного бустинга и нейронных сетей настоящей работы позволят точнее спрогнозировать не менее важные показатели, такие как ВВП, экспорт, обменные курсы, что будет способствовать проведению более эффективной макроэкономической политики в России.

*Приложение к статье см. на сайте*  
**<https://rjmf.econs.online>**

## Список литературы

- Абрамов В., Морозов А., Синяков А., Стерхова А.** О роли глобальных факторов в инфляции: аналитическая записка. – Москва: Банк России, 2022.
- Андреев А.** Исследование асимметрии и нелинейности переноса динамики обменного курса в инфляцию // Банк России. Серия докладов об экономических исследованиях. – 2019. – № 45.
- Андреев А.** Прогнозирование инфляции методом комбинирования прогнозов в Банке России // Банк России. Серия докладов об экономических исследованиях. – 2016. – № 14.
- Гафаров Б.** Кривая Филлипса и становление рынка труда в России // Экономический журнал ВШЭ. – 2011. – № 2. – С. 155–176.
- Доклад о денежно-кредитной политике.** – Москва: Банк России, 2019. – № 4(28).
- Кудрин А. Л.** Инфляция: российские и мировые тенденции // Вопросы экономики. – 2007. – № 10. – С. 4–26. doi: 10.32609/0042-8736-2007-10-4-26
- Орлов Д., Постников Е.** Кривая Филлипса: инфляция и NAIRU в российских регионах // Банк России. Серия докладов об экономических исследованиях. – 2020.
- Перевышин Ю. Н.** Краткосрочное прогнозирование инфляции в российской экономике // Экономическая политика. – 2022. – № 5. – С. 8–25. doi: 10.18288/1994-5124-2022-5-8-25
- Пономарев Ю. Ю., Трунин П. В., Улюкаев А. В.** Эффект переноса динамики обменного курса на цены в России // Вопросы экономики. – 2014. – № 3. – С. 21–35. doi: 10.32609/0042-8736-2014-3-21-35
- Синяков А. А., Чернядьев Д. Н., Сапова А. К.** Оценка эффекта переноса валютного курса рубля в цены производителей конечной продукции на микроданных российских компаний // Журнал Новой экономической ассоциации. – 2019. – № 1. – С. 128–157.
- Третьяков Д., Фокин Н.** Помогают ли высокочастотные данные в прогнозировании российской инфляции // Вестник СПбГУ. Экономика. – 2021. – № 37(2). – С. 318–343.
- Хабибуллин Р.** Какие показатели разрывов выпуска и реальной деловой активности позволяют прогнозировать инфляцию в России? // Банк России. Серия докладов об экономических исследованиях. – 2019. – № 50.
- Aanes B., Gullien M.** Forecasting Norwegian Inflation with Deep Neural Networks: Master Thesis / Norwegian School of Economics. – Bergen, 2018. URL: <https://openaccess.nhh.no/nhh-xmlui/bitstream/handle/11250/2586572/masterthesis.PDF> (дата обращения: 01.12.2023).

- Almosova A., Andresen N.** Nonlinear Inflation Forecasting with Recurrent Neural Networks // Journal of Forecasting. – 2023. – Vol. 42(2). – pp. 240–259. doi: 10.1002/for.2901
- Angelico C., Marcucci J., Miccoli M., Quarta F.** Can We Measure Inflation Expectations Using Twitter? // Journal of Econometrics. – 2022. – Vol. 228(2). – pp. 259–277. doi: 10.1016/j.jeconom.2021.12.008
- Araujo G., Gaglianone W.** Machine Learning Methods for Inflation Forecasting in Brazil: New Contenders Versus Classical Models // Latin American Journal of Central Banking. – 2023. – Vol. 4(2). – Article 100087. doi: 10.1016/j.latacb.2023.100087
- Athey S., Imbens G.** Machine Learning Methods that Economists Should Know About // Annual Review of Economics. – 2019. – Vol. 11. – pp. 685–725. doi: 10.1146/annurev-economics-080217-053433
- Atkinson A., Ohanian L.** Are Phillips Curve Useful for Forecasting Inflation? // Federal Reserve Bank of Minneapolis Quarterly Review. – 2001. – Vol. 25(1). – pp. 2–11. doi: 10.21034/qtr.2511
- Banbura M., Bobeica E.** Does the Phillips Curve help to Forecast Euro Area Inflation? // International Journal of Forecasting. – 2023. – Vol. 39(1). – pp. 364–390. doi: 10.1016/j.ijforecast.2021.12.001
- Barkan O., Benchimol J., Caspi I., Cohen E., Hammer A., Koenigstein N.** Forecasting CPI Inflation Components with Hierarchical Recurrent Neural Networks // International Journal of Forecasting. – 2023. – Vol. 39(3). – pp. 1145–1162. doi: 10.1016/j.ijforecast.2022.04.009
- Baybuza I.** Inflation Forecasting Using Machine Learning Methods // Russian Journal of Money and Finance. – 2018. – Vol. 77(4). – pp. 42–59. doi: 10.31477/rjmf.201804.42
- Breiman L.** Bagging Predictors // Machine Learning. – 1996. – Vol. 24. – pp. 123–140. doi: 10.1007/BF00058655
- Brownlee J.** Deep Learning for Time Series Forecasting. Predict the Future with MLPs, CNNs and LSTMs in Python. – Machine Learning Mastery, 2018.
- Chakraborty C., Joseph A.** Machine Learning at Central Banks // Bank of England Staff Working Paper. – 2017. – N 674.
- Chollet F.** Deep Learning with Python. – Shelter Island: Manning Publications Co., 2021.
- Chung J., Gulcehre C., Cho K., Bengio Y.** Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling // arXiv Preprint. – 2014. – arXiv:1412.3555. doi: 10.48550/arXiv.1412.3555
- Cochrane J.** Long-Term Debt and Optimal Policy in the Fiscal Theory of the Price Level // Econometrica. – 2001. – Vol. 69(1). – pp. 69–116. doi: 10.1111/1468-0262.00179
- Coibion O., Gorodnichenko Y., Kamdar R.** The Formation of Expectations, Inflation, and the Phillips Curve // Journal of Economic Literature. – 2018. – Vol. 56(4). – pp. 1447–1491. doi: 10.1257/jel.20171300
- Coulombe P. G.** A Neural Phillips Curve and a Deep Output Gap // arXiv Preprint. – 2022. – arXiv:2202.04146. doi: 10.48550/arXiv.2202.04146
- Coulombe P., Leroux M., Stevanovic D., Surprenant S.** How Is Machine Learning Useful for Macroeconomic Forecasting? // Journal of Applied Econometrics. – 2022. – Vol. 37(5). – pp. 920–964. doi: 10.1002/jae.2910

- Ditzen J., Ravazzolo F.** Dominant Drivers of National Inflation // arXiv Preprint. – 2022. – arXiv:2212.05841. doi: 10.48550/arXiv.2212.05841
- Dorogush A. V., Ershov V., Gulin A.** CatBoost: Gradient Boosting with Categorical Features Support // arXiv Preprint. – 2018. – arXiv:1810.11363. doi: 10.48550/arXiv.1810.11363
- Estrella A.** Why Does the Yield Curve Predict Output and Inflation? // The Economic Journal. – 2005. – Vol. 115(505). – pp. 722–744. doi: 10.1111/j.1468-0297.2005.01017.x
- Faust J., Wright J.** Forecasting Inflation // Handbook of Economic Forecasting, Vol. 2, Part A / Elliott G., Timmermann A., eds. – Elsevier, 2013. – pp. 2–56. doi: 10.1016/B978-0-444-53683-9.00001-3
- Fisher I.** I Discovered the Phillips Curve: ‘A Statistical Relationship Between Unemployment and Price Changes’ // Journal of Political Economy. – 1973. – Vol. 81(2). – Part 1. – pp. 496–502. doi: 10.1086/260048
- Forni M., Hallin M., Lippi M., Reichlin L.** Do Financial Variables Help Forecasting Inflation and Real Activity in the Euro Area? // Journal of Monetary Economics. – 2003. – Vol. 50(6). – pp. 1243–1255. doi: 10.1016/S0304-3932(03)00079-5
- Freund Y., Schapire R.** A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting // Journal of Computer and System Sciences. – 1997. – Vol. 55(1). – pp. 119–139. doi: 10.1006/jcss.1997.1504
- Friedman J.** Greedy Function Approximation: A Gradient Boosting Machine // The Annals of Statistics. – 2001. – Vol. 29(5). – pp. 1189–1232. doi: 10.1214/aos/1013203451
- Friedman M.** The Role of Monetary Policy // American Economic Review. – 1968. – Vol. 58(1). – pp. 1–17.
- Garcia M., Medeiros M., Vanconcelos G.** Real-Time Inflation Forecasting with High-Dimensional Models: The Case of Brazil // International Journal of Forecasting. – 2017. – Vol. 33(3). – pp. 679–693. doi: 10.1016/j.ijforecast.2017.02.002
- Garratt A., Lee K., Pesaran H., Shin Y.** Forecast Uncertainties in Macroeconomic Modeling: An Application to the U.K. Economy // Journal of the American Statistical Association. – 2003. – Vol. 98(464). – pp. 829–838.
- Geerolf F.** The Phillips Curve: A Relation Between Real Exchange Rate Growth and Unemployment. – 2020. URL: <https://fgeerolf.com/phillips.pdf> (дата обращения: 01.12.2023).
- Harding M., Lindé J., Trabandt M.** Understanding Post-COVID Inflation Dynamics // Journal of Monetary Economics. – 2023. – Vol. 140 (Supplement). – pp. S101–S118. doi: 10.1016/j.jmoneco.2023.05.012
- Hochreiter S., Schmidhuber J.** Long Short-Term Memory // Neural Computation. – 1997. – Vol. 9(8). – pp. 1735–1780. doi: 10.1162/neco.1997.9.8.1735
- Inoue A., Kilian L.** How Useful Is Bagging in Forecasting Economic Time Series? A Case Study of U.S. Consumer Price Inflation // Journal of the American Statistical Association. – 2008. – Vol. 103(482). – pp. 511–522. doi: 10.1198/016214507000000473
- Joseph A., Potjagailo G., Kalamara E., Chakraborty C., Kapetanios G.** Forecasting UK Inflation Bottom Up // Bank of England Staff Working Paper. – 2021. – N 915.
- Kapetanios G., Labhard V., Price S.** Forecasting Using Bayesian and Information-Theoretic Model Averaging: An Application to U.K. Inflation // Journal of Business and Economic Statistics. – 2008. – Vol. 26(1). – pp. 33–41.

- Karpathy A., Johnson J., Fei-Fei L.** Visualizing and Understanding Recurrent Networks // arXiv Preprint. – 2015. – arXiv:1506.02078. doi: 10.48550/arXiv.1506.02078
- Ke G., 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, Vol. 30 (Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, USA) / Guyon I., von Luxburg U., Bengio S., Wallach H., Fergus R., Vishwanathan S., Garnett R., eds. – 2017. – pp. 3149–3157.
- Kingma D., Ba J.** Adam: A Method for Stochastic Optimization // arXiv Preprint. – 2017. – arXiv:1412.6980. doi: 10.48550/arXiv.1412.6980
- Kinlaw W., Kritzman M., Metcalfe M., Turkington D.** The Determinants of Inflation // Journal of Investment Management. – 2023. – 21(3). – pp. 29–41.
- Kiselev A., Zhivaykina A.** The Role of Global Relative Price Changes in International Comovement of Inflation // Journal of Economic Asymmetries. – 2020. – Vol. 22. – Article e00175. doi: 10.1016/j.jeca.2020.e00175
- Kohlscheen E.** What Does Machine Learning Says about the Drivers of Inflation? // BIS Working Papers. – 2022. – N 980.
- Koop G., Korobilis D.** Forecasting Inflation Using Dynamic Model Averaging // International Economic Review. – 2012. – Vol. 53(3). – pp. 867–886. doi: 10.1111/j.1468-2354.2012.00704.x
- Koop G., Potter S.** Estimation and Forecasting in Models with Multiple Breaks // Review of Economic Studies. – 2007. – Vol. 74(3). – pp. 763–789. doi: 10.1111/j.1467-937X.2007.00436.x
- Longo L., Soltanieh-ha M.** SHAPoly: A Novel Shapley-Polynomial Framework for Estimating Nonlinear Dynamics in Macroeconomic Data Using Deep Neural Networks // SSRN Electronic Journal. – 2023. doi: 10.2139/ssrn.4350978
- Lundberg S., Lee S.-I.** A Unified Approach to Interpreting Model Predictions // arXiv Preprint. – 2017. – arXiv:1705.07874. doi: 10.48550/arXiv.1705.07874
- Maheu J., Gordon S.** Learning, Forecasting and Structural Breaks // Journal of Applied Econometrics. – 2008. – Vol. 23(5). – pp. 553–583. doi: 10.1002/jae.1018.
- Mamedli M., Shibitov D.** Forecasting Russian CPI with Data Vintages and Machine Learning Techniques // Bank of Russia Working Paper Series. – 2021.
- Masini R., Medeiros M., Mendes E.** Machine Learning Advances for Time Series Forecasting // Journal of Economic Surveys. – 2023. – Vol. 37(1). – pp. 76–111. doi: 10.1111/joes.12429
- Medeiros M., Schutte E., Soussi T.** Global Inflation: Implications for Forecasting and Monetary Policy // SSRN Electronic Journal. – 2023. doi: 10.2139/ssrn.4145665
- Medeiros M., Vasconcelos G., Veiga A., Zilberman E.** Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods // Journal of Business and Economic Statistics. – 2019. – Vol. 39(1). – pp. 98–119. doi: 10.1080/07350015.2019.1637745
- Nagy E., Tengey V.** External and Domestic Drivers of Inflation: The Case Study of Hungary. Russian Journal of Money and Finance. – 2018. – Vol. 77(3). – pp. 49–64. doi: 10.31477/rjmf.201803.49
- Nakamura E.** Inflation Forecasting Using a Neural Network // Economics Letters. – 2005. – Vol. 86(3). – pp. 373–378. doi: 10.1016/j.econlet.2004.09.003

- Paranhos L.** Predicting Inflation with Neural Networks // arXiv Preprint. – 2021. – arXiv:2104.03757. doi: 10.48550/arXiv.2104.03757
- Pavlov E.** Forecasting Inflation in Russia Using Neural Networks // Russian Journal of Money and Finance. – 2020. – Vol. 79(1). – pp. 57–73. doi: 10.31477/rjmf.202001.57
- Phelps E.** Money-Wage Dynamics and Labor-Market Equilibrium // Journal of Political Economy. – 1968. – 76(4). – pp. 678–711. doi: 10.1086/259438
- Phillips A.** The Relation Between Unemployment and the Rate of Change of Money Rates in the United Kingdom, 1861–1957 // Economica. – 1958. – Vol. 25(100). – pp. 283–299. doi: 10.2307/2550759
- Prokhorenkova L., Gusev G., Vorobev A., Dorogush A. V., Gulín A.** CatBoost: Unbiased Boosting with Categorical Features // Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18). – Curran Associates Inc., 2018. – pp. 6639–6649.
- Saul S.** Do Global Output Gaps Help Forecast Inflation in Russia? // Bank of Russia Working Paper Series. – 2021. – N 85.
- Semiturkin O., Shevelev A.** Correct Comparison of Predictive Features of Machine Learning Methods: The Case of Forecasting Inflation Rates in Siberia // Russian Journal of Money and Finance. – 2023. – Vol. 82(1). – pp. 87–103.
- Shulyak E.** Macroeconomic Forecasting Using Data from Social Media // Russian Journal of Money and Finance. – 2022. – Vol. 81(4). – pp. 86–112.
- Sims C.** Stepping on a Rake: The Role of Fiscal Policy in the Inflation of the 1970s // European Economic Review. – 2010. – Vol. 55(1). – pp. 48–56. doi: 10.1016/j.eurocorev.2010.11.010
- Stella A., Stock J.** A State-Dependent Model for Inflation Forecasting // Board of Governors of the Federal Reserve System International Finance Discussion Papers. – 2012. – N 1062.
- Stock J., Watson M.** Forecasting Inflation // Journal of Monetary Economics. – 1999. – Vol. 44(2). – pp. 293–335. doi: 10.1016/S0304-3932(99)00027-6
- Stock J., Watson M.** Forecasting Output and Inflation: The Role of Asset Prices // Journal of Economic Literature. – 2003 – Vol. 41(3). – pp. 788–829. doi: 10.1257/002205103322436197
- Stock J., Watson M.** Modeling Inflation After the Crisis // NBER Working Paper. – 2010. – N 16488.
- Stock J., Watson M.** Phillips Curve Inflation Forecasts // NBER Working Paper. – 2008. – N 14322.
- Stock J., Watson M.** Why Has U.S. Inflation Become Harder to Forecast? // Journal of Money, Credit and Banking. – 2007. – Vol. 39(s1). – pp. 3–33. doi: 10.1111/j.1538-4616.2007.00014.x
- Styrin K.** Forecasting Inflation in Russia by Dynamic Model Averaging // Russian Journal of Money and Finance. – 2019. – Vol. 78(1). – pp. 3–18. doi: 10.31477/rjmf.201901.03
- Styrin K., Zamulin O.** A Real Exchange Rate Based Phillips Curve // CEFIR / NES Working Paper Series. – 2012. – N 179.
- Svensson L.** Inflation Targeting // Handbook of Monetary Economics, Vol. 3 / Friedman B., Woodford M., eds. – Elsevier, 2010. – pp. 1237–1302.
- Szafranek K.** Bagged Neural Networks for Forecasting Polish (Low) Inflation // International Journal of Forecasting. – 2019. – Vol. 35. – pp. 1042–1059. doi: 10.1016/j.ijforecast.2019.04.007

- 
- Van den Oord A., Dieleman S., Zen H., Simonyan K., Vinyals O., Graves A., Kalchbrenner N., Senior A., Kavukcuoglu K.** WaveNet: A Generative Model for Raw Audio // arXiv Preprint. – 2016. – arXiv:1609.03499. doi: 10.48550/arXiv.1609.03499
- Zeng J.** Forecasting Aggregates with Disaggregate Variables: Does Boosting Help to Select the Most Relevant Predictors? // Journal of Forecasting. – 2017. – Vol. 36(1). – pp. 74–90. doi: 10.1002/for.2415