

# Алгоритмы машинного обучения

для краткосрочного  
прогнозирования темпов  
прироста реального ВВП

Цукарев Т., Погосян К., Лемба К.

**РД/24/5**

**Ключевые слова:** краткосрочное прогнозирование, методы машинного обучения, темп прироста реального ВВП, прогноз, макроэкономические показатели, взвешенный прогноз.

**Коды JEL:** C11, C13, C52, C53.

**Контакты авторов:**

Тарас Цукарев, начальник отдела макроэкономического анализа и статистики Группы главного экономиста ЕФСР: [ttsukarev@efsd.org](mailto:ttsukarev@efsd.org)

Карен Погосян, старший экономист отдела макроэкономического анализа и статистики Группы главного экономиста ЕФСР: [kpoghosyan@efsd.org](mailto:kpoghosyan@efsd.org)

Кирилл Лемба, эксперт отдела макроэкономического анализа и статистики Группы главного экономиста ЕФСР: [KLemba@efsd.org](mailto:KLemba@efsd.org)

Электронная версия документа находится в свободном доступе по электронному адресу: <https://efsd.org/research/working-papers/>

Рабочие документы ЕФСР представляют собой аналитические исследования в области глобальных, региональных и страновых экономических тенденций, экономического моделирования, макроэкономического анализа, отраслевого анализа, международной финансовой архитектуры и иных вопросов. Выводы и интерпретация факторов в рабочем документе не обязательно отражают официальную позицию Экспертного совета ЕФСР, Совета ЕФСР или государств — учредителей ЕФСР.

Перепечатка текста в некоммерческих целях, целиком или по частям, включая крупные фрагменты, и размещение текста на внешних электронных ресурсах разрешены при обязательной ссылке на оригинальный текст.

# Оглавление

Список сокращений .....	2
Благодарность .....	3
Введение .....	4
1. Обзор моделей .....	7
2. Данные .....	17
3. Результаты экспериментов .....	18
Заключение .....	24
Список литературы .....	25
Приложение .....	27

## Список сокращений

<b>Белстат</b>	Национальный статистический комитет Республики Беларусь
<b>ВВП</b>	валовой внутренний продукт
<b>ЕФСР</b>	Евразийский фонд стабилизации и развития
<b>МБК</b>	межбанковский кредит
<b>МНК</b>	метод наименьших квадратов
<b>НБРБ</b>	Национальный банк Республики Беларусь
<b>RSS</b>	сумма квадратов остатков (residuals sum of squares)
<b>доллар</b>	доллар США
<b>млн</b>	миллион

## Благодарность

Авторы выражают благодарность Татевик Сехпосян (доценту экономического факультета, Texas A&M University), Александру Новопольцеву (советнику Управления исследований и стратегического развития, Национальный банк Республики Беларусь), а также сотрудникам ЕФСР за ценные комментарии и рекомендации. Ответственность за любые ошибки и недочеты остается за авторами.

## Введение

Высокая точность краткосрочного прогноза изменения экономической активности имеет существенное значение для принятия своевременных решений в области макроэкономической политики. Однако достижение данной точности может осложняться рядом причин. Во-первых, часть важных макроэкономических показателей формируется статистическими ведомствами на квартальной, а не месячной основе, что ограничивает оперативность анализа. Во-вторых, публикация официальной статистической информации нередко осуществляется с определенным запаздыванием относительно текущих процессов в экономике. В-третьих, для анализа и прогнозирования макроэкономической динамики используются, как правило, квартальные структурные или полуструктурные модели, которые имеют недостаток в виде не всегда качественной оценки тенденций в краткосрочной перспективе (1–2 квартала). При этом от начальной точки во многом зависит результат среднесрочного прогноза (Жемков, 2021). Таким образом, существует объективная необходимость в точной и оперативной оценке макроэкономических переменных, в том числе ВВП.

Международная практика показывает, что в арсенале у исследователей и экспертов, которые занимаются краткосрочным прогнозированием ВВП, имеется широкий перечень инструментов, как правило, статистических и эконометрических методов и моделей. В последние годы у экономистов все большую популярность приобретают методы (алгоритмы) машинного обучения, которые выступили своеобразной альтернативой традиционным эконометрическим моделям (Desai, 2023).

Методы и алгоритмы машинного обучения имеют ряд значительных преимуществ по сравнению с традиционным модельным инструментарием, применяемым для анализа и прогнозирования макроэкономических показателей. Во-первых, они способны работать с большими массивами данных, когда количество регрессоров (предикторов) больше, чем количество наблюдений. Во-вторых, алгоритмы машинного обучения могут применяться как для решения задачи регрессии<sup>1</sup>, так и для задачи классификации. В-третьих, в отличие от стандартных эконометрических моделей, алгоритмы машинного обучения позволяют работать с несбалансированными данными (*imbalanced datasets*), используя при этом всю доступную информацию. Еще один важный момент, на который часто не обращают внимания, касается линейности связей. Если стандартные эконометрические модели требуют от эконометриста явного контроля за нелинейностью и взаимосвязями, об их существовании он должен знать или предполагать априори, методы машинного обучения разработаны таким образом, чтобы самостоятельно учитывать подобного рода взаимосвязи и динамику. Обозначенные преимущества способствуют тому, что методы (алгоритмы) машинного обучения зачастую обладают более высокой прогностической эффективностью (Alessi and Savona, 2021).

Несмотря на ряд преимуществ, методы и алгоритмы машинного обучения имеют несколько ограничений и недостатков (Desai, 2023):

- для обучения моделей, особенно моделей глубокого машинного обучения, часто требуются большие массивы данных и значительные вычислительные ресурсы;

---

<sup>1</sup> В данной работе акцентируется внимание на применении алгоритмов машинного обучения для целей регрессии.

- из-за своей гибкости модели могут легко переобучиться<sup>2</sup>, в результате вероятность ошибки обученного алгоритма на объектах тестовой выборки существенно выше, чем средняя ошибка на обучающей выборке. Переобучение может возникнуть при использовании чрезмерно сложных моделей;
- методы и алгоритмы машинного обучения могут иметь смещенную оценку прогноза (проблема выбора между смещением и дисперсией (*bias and variance tradeoff*)), если, например, данные для обучения моделей имеют низкое качество. В результате большого смещения алгоритм может пропустить связь между признаками и выводом;
- сложность вычислительных процессов затрудняет интерпретацию полученных результатов (проблема «черного ящика»).

Однако исследователи и аналитики данных добиваются прогресса в преодолении ограничений для моделей машинного обучения, все чаще используя их в своей работе. Важно отметить, что машинное обучение может дополнять традиционные статистические и эконометрические инструменты, выявляя сложные взаимосвязи и закономерности в данных, которые затем могут быть включены в эконометрические модели. Поскольку экономика и экономические данные продолжают усложняться, методы машинного обучения остаются ценным инструментом для экономического анализа и прогнозирования.

Цель настоящего исследования — оценить, могут ли методы и алгоритмы машинного обучения потенциально повысить точность прогнозов и тем самым стать полезным дополнением к традиционному инструментарию для краткосрочного прогнозирования либо полезной альтернативой.

Данный рабочий документ дополняет растущее количество исследований, посвященных применению методов и алгоритмов машинного обучения для макроэкономического прогнозирования. Например, в работе [Buckmann et al. \(2021\)](#) представлено комплексное сравнительное исследование использования моделей машинного обучения для макроэкономического прогнозирования. Авторами было установлено, что модели машинного обучения в основном превосходят традиционные эконометрические подходы при прогнозировании изменения уровня безработицы в США на горизонте до одного года. [Chakraborty and Joseph \(2017\)](#) исследовали области применения моделей машинного обучения центральными банками и пришли к выводу, что этот вид моделей имеет высокий потенциал в области поддержки принятия решений и подготовки макроэкономических прогнозов. Например, [Biau and D’Elia \(2010\)](#), используя алгоритм Random Forest для прогнозирования ВВП еврозоны, установили, что данный алгоритм машинного обучения способен превзойти по точности прогнозы, полученные с помощью стандартной AR-модели. В работе [Tiffin \(2016\)](#) продемонстрировано, что, несмотря на двухгодичный лаг публикации официальной статистики по ВВП Ливана, алгоритмы Elastic Net и Random Forest позволяют устранить эту проблему и могут являться достойными инструментами для определения своевременной оценки экономического роста в развивающихся странах. [Tkacz and Hu \(1999\)](#), в свою очередь, предлагают для прогнозирования темпов

---

<sup>2</sup> Переобучение (англ. *overfitting*) — негативное явление, возникающее в случае, когда алгоритм обучения модели реализует прогноз, который слишком близко или точно соответствует конкретному набору данных и поэтому не подходит для применения алгоритма к дополнительным данным или новым наблюдениям.

прироста ВВП использовать искусственные нейронные сети (*artificial neural networks*). При помощи нейронных сетей авторам удалось получить прогнозы на 15–19% точнее, чем с использованием традиционных линейных моделей. Аналогичным образом [Chuku, Oduor и Simpasa \(2019\)](#) на примере временных рядов ряда африканских стран продемонстрировали, что алгоритмы нейронных сетей дают заметно лучшую точность прогноза, чем традиционные структурные эконометрические модели и ARIMA-модели.

В данной работе рассматриваются три традиционные (стандартные) модели для краткосрочного прогнозирования — AR, VAR и BVAR. В качестве сравнительной альтернативы к ним взяты такие методы и алгоритмы машинного обучения, как Ridge- и LASSO-регрессии, Elastic Net, Boosting, Bagging, Random Forest и Neural Networks.

Для оценивания и прогнозирования по традиционным моделям и алгоритмам машинного обучения использовались фактические квартальные макроэкономические временные ряды показателей реального, внешнего, денежно-кредитного и финансового секторов экономики Республики Беларусь для периода 2004–2024 годов. Часть временных рядов была подвергнута первоначальной статистической обработке, а именно сезонному сглаживанию, логарифмированию и взятию первых разностей.

С целью определения наиболее эффективного с точки зрения прогностических свойств метода прогнозирования проведены различные эксперименты. Их суть в следующем. Общее количество наблюдений во временных рядах делится на две части. Далее первая часть фактических данных применяется для оценивания параметров моделей, а вторая часть наблюдений — для сравнения полученных прогнозов по традиционным моделям и по методам машинного обучения с фактическими данными. Генерация прогнозов по различным методам и моделям осуществляется по рекурсивной схеме регрессионного эксперимента ([Poghosyan and Magnus, 2012](#)). Наилучшим считается тот инструмент, который минимизирует выбранную для этих целей функцию потерь.

Результаты наших экспериментов показывают, что прогнозы, полученные с помощью алгоритмов машинного обучения, по точности превосходят прогнозы традиционных моделей. Наиболее эффективными инструментами оказались LASSO-регрессия, Random Forest и Neural Networks. С другой стороны, согласно статистике Диболда — Мариано ([Diebold and Mariano, 1995](#)), существенных различий в полученных прогнозах по традиционным моделям и алгоритмам машинного наблюдения не наблюдалось. При этом расчеты продемонстрировали, что комбинирование<sup>3</sup> различных прогнозов позволяет еще больше повысить точность прогнозирования.

Данное исследование имеет следующую структуру. В первом разделе приведен краткий обзор применяемых в работе моделей и алгоритмов машинного обучения. Во втором разделе представлен обзор анализируемых данных и используемых методов предварительной статистической обработки временных рядов. В третьем разделе описываются эксперименты и полученные на их основе результаты. Заключение содержит основные выводы и рекомендации.

---

<sup>3</sup> В частности, путем взвешивания на основе МНК весов.

# 1. Обзор моделей

Приведем описание сравниваемых в настоящей работе традиционных эконометрических моделей, таких как авторегрессия, векторная авторегрессия и байесовская векторная авторегрессия (в дальнейшем AR, VAR и BVAR соответственно), а также методов и алгоритмов машинного обучения.

## Традиционные эконометрические модели

### Модель одномерной авторегрессии (AR)

AR-модель можно представить следующим образом (Richardson et al., 2018):

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t, \quad (1)$$

где  $\phi_0, \phi_1$  — параметры;  
 $\varepsilon_t$  — случайные ошибки модели.

### Модель векторной авторегрессии (VAR)

Традиционная VAR-модель без ограничений имеет вид (Pirschel and Wolters, 2014):

$$Y_t = c + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + \varepsilon_t, \quad (2)$$

где  $Y_t$  — вектор исходных переменных размера  $(n \times 1)$ ;  
 $c$  — вектор постоянных коэффициентов размера  $(n \times 1)$ ;  
 $A_1, \dots, A_p$  — матрицы параметров размера  $(n \times n)$  для различных временных лагов  $l = 1, 2, \dots, p$ ;  
 $\varepsilon_t$  — вектор случайных ошибок модели размера  $(n \times 1)$ ,  $\varepsilon_t \sim N(0, \sigma^2 I_{n \times n})$ .

Неизвестные параметры векторной авторегрессии можно оценить с помощью метода наименьших квадратов (МНК) (Hamilton, 1994).

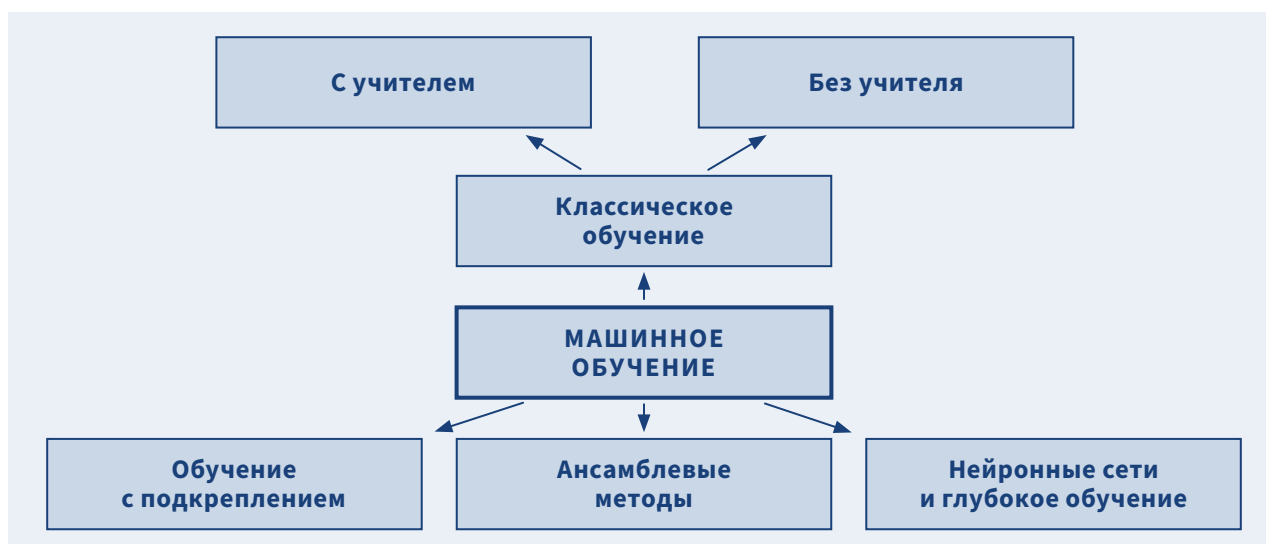
### Модель байесовской векторной авторегрессии (BVAR)

Основная идея BVAR-модели состоит в следующем: при увеличении временного лага объясняющих переменных оцениваемые параметры обнуляются, что позволяет решить проблему сверхпараметризации модели. В этом отношении, по сравнению с традиционной VAR, BVAR является моделью регуляризации (Gupta and Kabundi, 2009). В соответствии с BVAR штраф на параметры модели накладывается исходя из априорной информации о нормальной природе распределения параметров с нулевым математическим ожиданием и убывающим стандартным отклонением (Poghosyan, 2015). Исключение — параметр при первом лаге зависимой переменной, у которого математическое ожидание приравнивается к единице (Litterman, 1980; Hamilton, 1994).

## Методы и алгоритмы машинного обучения

Прежде чем представить многообразие методов и алгоритмов машинного обучения, обратимся к известной публикации американского математика-статистика Лео Бреймана, где он сделал заключение, что существуют две школы или «культуры» статистического моделирования данных (Breiman, 2001). Одна из них предполагает наличие определенной стохастической модели, лежащей в основе процесса генерации данных. Другая — скорее пытается найти функцию, которая наилучшим образом предсказывает результаты для определенных входных данных. Ко второй «культуре» как раз и относится алгоритмическое моделирование, или машинное обучение. Общая черта подходов к машинному обучению в том, что алгоритмы реализуются с минимальным вмешательством человека. Один из вариантов таксономии, используемой для классификации алгоритмов машинного обучения, представлен на [Рисунке 1](#).

**Рисунок 1. Основные виды машинного обучения**



**Источник:** разработка авторов.

Классические алгоритмы машинного<sup>4</sup> обучения начали разрабатывать еще в 1950-е годы, это одни из наиболее популярных инструментов для прогнозирования за счет своей относительной простоты. Классическое обучение, как правило, разделяют на два вида: с учителем (*supervised learning*) и без учителя (*unsupervised learning*).

Алгоритм обучения с учителем берет известный набор входных данных и известную реакцию на эти данные (выход) и обучает модель генерировать адекватный прогноз реакции на новые данные. К обучению с учителем относятся алгоритмы на основе регрессий (Ridge- и LASSO-регрессии, Elastic Net и др.) и классификаций (дерева решений, метод опорных векторов, метод К-ближайших соседей и др.).

<sup>4</sup> Классическое машинное обучение строится на классических статистических алгоритмах, таких как линейная регрессия, логистическая регрессия, дерево решений и т. п. Глубокое машинное обучение, в отличие от классического, имеет достаточно сложную архитектуру алгоритмов решения задач. Кроме того, в отличие от классических алгоритмов машинного обучения, глубокое обучение требует огромной вычислительной мощности.

Обучение без учителя позволяет находить в данных скрытые закономерности или внутренние структуры. При этом процесс обучения модели не предполагает наличия ответа в данных (отсутствие разметки данных) — модель должна сама найти закономерности и кластеризовать данные. К обучению без учителя можно отнести следующие методы: кластеризация (метод К-средних, Mean-Shift, DBSCAN и др.), поиск ассоциативных правил (FP-growth, Apriori, Eclat), снижение размерности (метод главных компонент, сингулярное разложение, латентно-семантический анализ и др.) (Alessi and Savona, 2021).

*Ансамблевые методы* предполагают одновременное обучение нескольких моделей для решения одной задачи. Полученные на основе используемых моделей прогнозы объединяются для генерации более качественных результатов (Alessi and Savona, 2021). Выделяют, как правило, следующие ансамблевые методы:

- стекинг (Stacking),
- бэггинг (Bagging, Bootstrap Aggregating),
- метод случайного леса (Random Forest),
- бустинг (Boosting), включает такие алгоритмы градиентного бустинга, как CatBoost, XGBoost и LightGBM.

*Обучение с подкреплением (Reinforcement Learning)* — это метод машинного обучения, в котором модель обучается методом проб и ошибок. Модель решает задачу, взаимодействуя с окружающим миром, улучшая свое поведение на основе полученного в ходе взаимодействия опыта. К наиболее популярным алгоритмам обучения с подкреплением можно отнести Q-Learning, Deep Q-Networks, State-Action-Reward-State-Action (SARSA), генетический алгоритм (Vaish et al., 2021).

Первые теоретические разработки *нейронных сетей (глубокого обучения)* датируются концом 1950-х годов, тогда как их техническая реализация состоялась лишь в 2011 году. К настоящему времени нейросети становятся одним из лидеров в области машинного обучения и решают самые разнообразные задачи во многих сферах деятельности человека.

Нейронные сети — это семейство моделей, созданных для имитации механизмов обучения человеческого мозга. Другими словами, искусственные нейросети состоят из «нейронов», которые взаимодействуют между собой. Нейрон, получая входные (исходные) данные или сигналы (*input*), преобразует их и передает другим нейронам, что в конечном итоге приводит к выходу искомого результата (*output*). Разница между получаемым на выходе результатом и целью используется для корректировки весов, применяемых для соединения нейронов, что позволяет минимизировать ошибку. К наиболее популярным архитектурам нейронных сетей можно отнести многослойный перцептрон (Multilayer Perceptron, MLP), сверточные сети (Convolutional Neural Networks, CNN), рекуррентные сети (Recurrent Neural Networks, RNN), автоэнкодеры (Autoencoder, AE) (Barbaglia et al., 2021).

В текущем исследовании мы будем использовать наиболее популярные методы и алгоритмы машинного обучения (Ridge- и LASSO-регрессии, Elastic Net, Boosting, Bagging, Random Forest и Neural Networks).

## Классические алгоритмы машинного обучения с учителем

### Ridge-регрессия (гребневая регрессия)

Алгоритм гребневой регрессии схож с МНК. Он разработан для преодоления нестабильности оценок МНК<sup>5</sup> путем пенализации суммы квадратов параметров на основе регуляризации<sup>6</sup>  $L_2(\sum_{j=1}^p \beta_j^2)$ . В частности, коэффициенты модели оцениваются по следующей оптимизационной формуле (Hoerl and Kennard, 1970):

$$\hat{\beta} = \arg \min_{\beta} \left[ \underbrace{\sum_{i=1}^n (y_i - \sum_{j=1}^p x_{ij} \beta_j)^2}_{\text{RSS}} + \lambda \underbrace{\sum_{j=1}^p \beta_j^2}_{\text{Ridge-пенализация}} \right], \quad (3)$$

где  $n$  — количество наблюдений;  
 $p$  — количество объясняющих переменных;  
 $\lambda \geq 0$  — параметр, который контролирует величину сжатия<sup>7</sup>.

Перепишав функционал в матричном виде:

$$\text{RSS}(\lambda) = (Y - X\beta)^T (Y - X\beta) + \lambda \beta^T \beta. \quad (4)$$

Взяв производные по  $\beta$ , получим оценки неизвестных параметров:

$$\hat{\beta} = (X^T X + \lambda I)^{-1} X^T Y, \quad (5)$$

где  $I$  — единичная матрица.

Согласно формуле (5), в отличие от МНК, к диагонали матрицы  $X^T X$  добавляется положительная константа  $\lambda$ . Если  $\lambda = 0$ , то оценки Ridge-регрессии будут аналогичны результатам МНК, а если  $\lambda \rightarrow \infty$ , то коэффициенты модели стремятся к нулю. Таким образом, на практике параметр  $\lambda$  изменяется между 0 и  $\infty$ .

Оценить параметры Ridge-регрессии можно по вышеуказанной аналитической формуле. Для этого можно применять и алгоритм градиентного спуска (*gradient descent*), что позволит сэкономить много времени на расчетах. Сходимость данного метода зависит от выбранных начальных параметров.

### LASSO-регрессия (оператор наименьшего абсолютного сжатия и выбора)

В случае применения LASSO-регрессии для сжатия коэффициентов используется регуляризация  $L_1(\sum_{j=1}^p |\beta_j|)$ , что и является основным отличием от Ridge-регрессии (Tibshirani, 1996):

$$\hat{\beta} = \arg \min_{\beta} \left[ \underbrace{\sum_{i=1}^n (y_i - \sum_{j=1}^p x_{ij} \beta_j)^2}_{\text{RSS}} + \lambda \underbrace{\sum_{j=1}^p |\beta_j|}_{\text{LASSO-пенализация}} \right]. \quad (6)$$

<sup>5</sup> При высокой корреляции между независимыми переменными (мультиколлинеарности).

<sup>6</sup> Регуляризация — это способ уменьшить сложность модели с целью предотвращения ее переобучения. Регуляризация  $L_1$  штрафует весовые значения добавлением суммы их абсолютных значений к ошибке. Регуляризация  $L_2$  выполняет аналогичную операцию добавлением суммы их квадратов к ошибке.

<sup>7</sup> Чем больше значение  $\lambda$ , тем больше величина сжатия.

Ограничение  $\sum_{j=1}^p |\beta_j|$  делает решение нелинейным, в этой связи для решения применяют алгоритм градиентного спуска.

### Elastic Net (Эластичная сеть)

Elastic Net является взвешенной комбинацией Ridge- и LASSO-регрессий и имеет следующую оптимизационную формулу:

$$\hat{\beta} = \arg \min_{\beta} \left[ \underbrace{\sum_{i=1}^n (y_i - \sum_{j=1}^p x_{ij} \beta_j)^2}_{\text{RSS}} + \lambda \sum_{j=1}^p \left[ \underbrace{(1 - \alpha) \beta_j^2}_{\text{Ridge}} + \underbrace{\alpha |\beta_j|}_{\text{LASSO}} \right] \right], \quad (7)$$

где  $\alpha$  — вес пенализации.

## Ансамблевые методы машинного обучения

Ансамблевые методы машинного обучения связаны с деревом решений, поэтому представим пояснения для дерева решений, прежде чем описывать используемые в данном исследовании методы.

Основная задача дерева решений — оптимальным образом разделить исходное множество наблюдений (из родительского узла) на два подмножества (или на два дочерних узла).

Пусть  $x_{nd}$  —  $d$ -й предиктор из вектора  $x_n$ . Если  $x_{nd}$  — количественная переменная, то правило разделения имеет вид:  $x_{nd} \leq t$ , где  $t$  — некоторый пороговый уровень.

Пусть  $N_m$  —  $m$ -й узел дерева решений, а  $C_m^L$  и  $C_m^R$  — два дочерних узла для  $N_m$ . Тогда наблюдения в узле  $N_m$  попадут в дочерний узел  $C_m^L$ , если  $x_{nd} \leq t$ , и в узел  $C_m^R$  — в обратном случае. Пусть  $n \in N_m$  — количество наблюдений в узле  $N_m$ , а  $\bar{y}_m$  — среднее значение узла  $N_m$ . Тогда сумма квадратов ошибок (RSS) рассчитывается следующим образом:

$$RSS_m = \sum_{n \in N_m} (y_n - \bar{y}_m)^2, \quad (8)$$

где  $y_n$  — наблюдения в узле  $N_m$ . Функция потерь для всего дерева решений — общий показатель  $RSS_T$ , который включает потери бутонов и листьев. Пусть  $I_m$  — индекс  $m$ -го бутона или листка. Тогда общая потеря всего дерева будет равна:

$$RSS_T = \sum_m \sum_{n \in N_m} I_m RSS_m. \quad (9)$$

Основным критерием оптимального разделения является минимизация показателя  $RSS_T$ . Его сокращение в результате разделения узла  $N_m$  на дочерние  $C_m^L$  и  $C_m^R$  рассчитывается следующим образом:

$$\Delta RSS_T = RSS_m - (RSS_{C_m^L} + RSS_{C_m^R}). \quad (10)$$

Как видим,  $\Delta RSS_T$  минимизируется, когда различие между двумя дочерними узлами  $RSS_{C_m^L}$  и  $RSS_{C_m^R}$  максимизируется (Breiman and Ihaka, 1984).

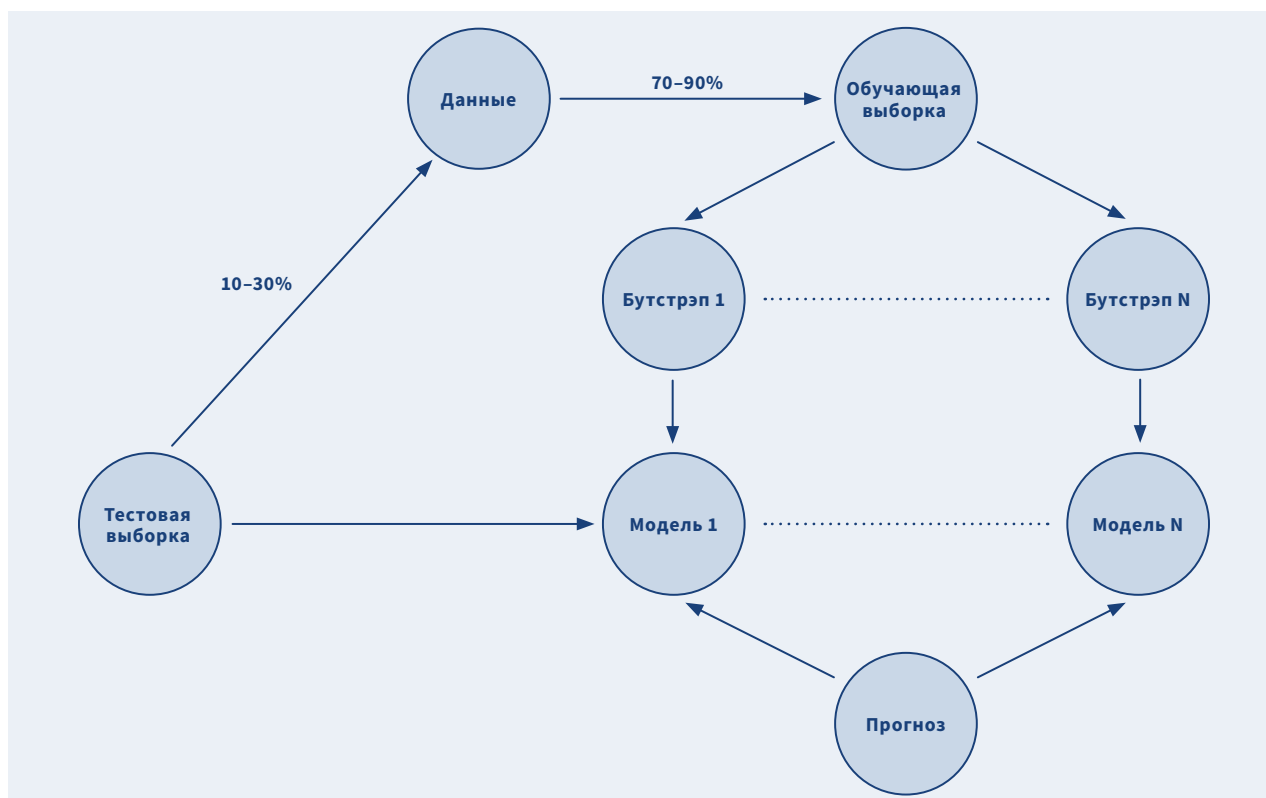
## Bagging (бэггинг)

Бэггинг, или бутстрэп-агрегирование, основывается на дереве решений и включает следующие шаги (Breiman, 1996; Hastie et al., 2009):

1. Задается общее количество бутстрэп-повторений<sup>8</sup>, максимальная и минимальная глубина дерева решений<sup>9</sup> (Efron and Tibshirani, 1993).
2. Для каждой бутстрэп-выборки строится дерево решений и оцениваются средние величины зависимой переменной на концах дерева (листьях). Полученные результаты прогнозов по данной бутстрэп-выборке сохраняются.
3. На основе тестовой выборки прогнозируются величины зависимой переменной, результат сохраняется.
4. Указанные первые три шага повторяются для заданного количества бутстрэп-повторений. Полученные прогнозы зависимой переменной по всем бутстрэп-повторениям усредняются, что является окончательным прогнозом зависимой переменной.

Визуализация бутстрэп-агрегирования представлена на [Рисунке 2](#).

**Рисунок 2. Визуализация бутстрэп-агрегирования**



**Источник:** разработка авторов.

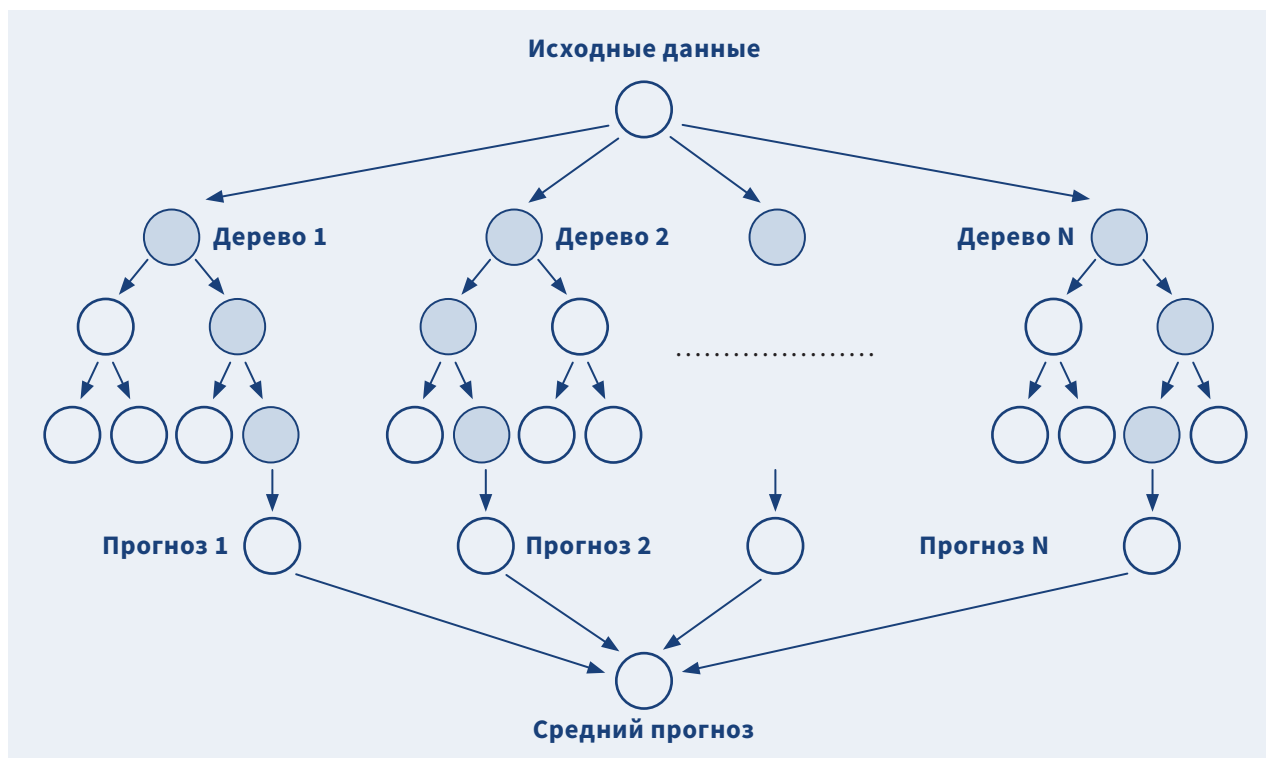
<sup>8</sup> Как правило, около 100 бутстрэп-повторений достаточно для получения стабильных прогнозов дерева решений.

<sup>9</sup> Максимальную глубину задают таким образом, чтобы не произошло переобучения дерева решений, а минимальный размер обычно равен двум.

### Random Forest (метод случайного леса)

Основное различие состоит в том, что если в бэггинге применяются все включенные в анализ объясняющие переменные, то в алгоритме Random Forest применяются лишь некоторые из случайно отобранных переменных (Breiman, 2001). Визуализация метода случайного леса представлена на Рисунке 3.

Рисунок 3. Визуализация метода случайного леса



Источник: разработка авторов.

### Boosting (бустинг)

Бустинг также основан на применении дерева решений и предполагает следующие шаги (Freund and Schapire, 1997):

1. Задаются исходные значения весов  $w_n^1$ , где  $n = 1, 2, \dots, N$ . Отметим, что  $N$  — это количество строк обучающей выборки.
2. Задается общее количество бутстрэп-повторений, например  $T = 100$ . Для каждого бутстрэп-повторения строится дерево решений и прогнозируются средние значения зависимой переменной на концах дерева (листьях) на основе обучающей выборки.
3. Далее рассчитывается вектор ошибок  $L_n^t$  для  $n = 1, 2, \dots, N$ :

$$L_n^t = \frac{|y_n - \hat{y}_n|}{D^t}, \quad (11)$$

где  $D^t = \max_n \{|y_n - \hat{y}_n|\}$ .

4. Рассчитываем среднее значение ошибок наблюдения  $\bar{L}^t$ :

$$\bar{L}^t = \sum_{n=1}^N L_n^t w_n^t. \quad (12)$$

Если  $\bar{L}^t \geq 0,5$ , то итерационный процесс завершаем, тогда общее количество бутстрэп-повторений  $T = 100$  заменяем на  $t - 1$ , где  $t = 1, 2, \dots, T$ . В обратном случае итерационный процесс продолжается, переходим к следующему шагу.

5. Рассчитываем  $\beta^t = \frac{\bar{L}^t}{1 - \bar{L}^t}$ . Чем меньше  $\beta^t$ , тем больше доверие к модели.

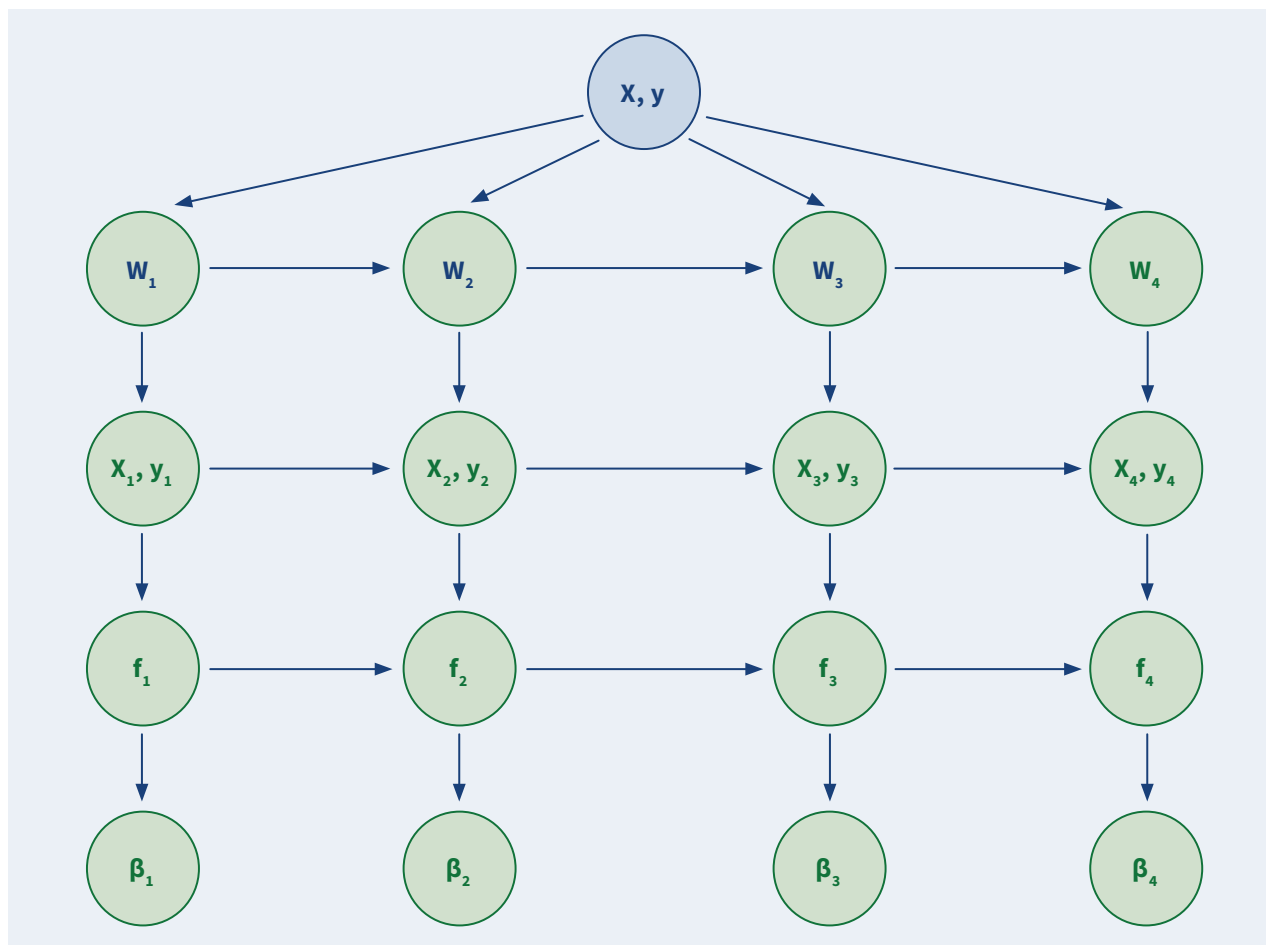
6. Обновляем весовые коэффициенты модели по следующему правилу:  $w_n^{t+1} = \frac{w_n^t (\beta^t)^{1-L_n^t}}{\sum_{n=1}^N w_n^t (\beta^t)^{1-L_n^t}}$ ,

что уменьшает вес наблюдений со сравнительно большей ошибкой.

7. Если  $\bar{L}^t \geq 0,5$ , рассчитываем значения прогнозов зависимой переменной как взвешенную медианную величину с помощью весов  $\ln\left(\frac{1}{\beta^t}\right)$ .

Визуализация бустинга представлена на [Рисунке 4](#).

**Рисунок 4. Визуализация бустинга**



Источник: разработка авторов.

## Neural Networks (нейронные сети)

Алгоритм построения нейронной сети, используемой для прогнозирования, состоит из нескольких этапов (Hastie et al., 2009):

1. Задается количество скрытых слоев (*hidden layers*). Пусть количество скрытых слоев равно  $n_{hidden} = 2$ .
2. Для каждого скрытого слоя и с учетом количества используемых предикторов случайным образом генерируются веса  $W_1$ . Так, размерность матрицы весов для предикторов будет  $(2 \times 17)$ , так как  $n_{hidden} = 2$ , а  $D_x = 17$ , где  $D_x$  — количество столбцов в матрице предикторов (пример визуализации рассматриваемой нейронной сети представлен на Рисунке 5). Процедура по части зависимой переменной такая же.
3. Генерируются матрицы весов для постоянной компоненты модели  $c_1$  размерностью  $(n_{hidden} \times 1)$ , а также для зависимой переменной  $W_2$ . Так как в модели одна зависимая переменная, то размерность сгенерированной матрицы будет  $(1 \times n_{hidden})$ . Соответственно, матрица для константы  $c_2$  будет  $(1 \times 1)$ .
4. На основе сгенерированных матриц весов рассчитываются выходные переменные:

$$h_1 = W_1 X' + c_1, \quad (13)$$

$$z_1 = f_1(h_1), \quad (14)$$

$$h_2 = W_2 z_1 + c_2, \quad (15)$$

$$\hat{y} = f_2(h_2), \quad (16)$$

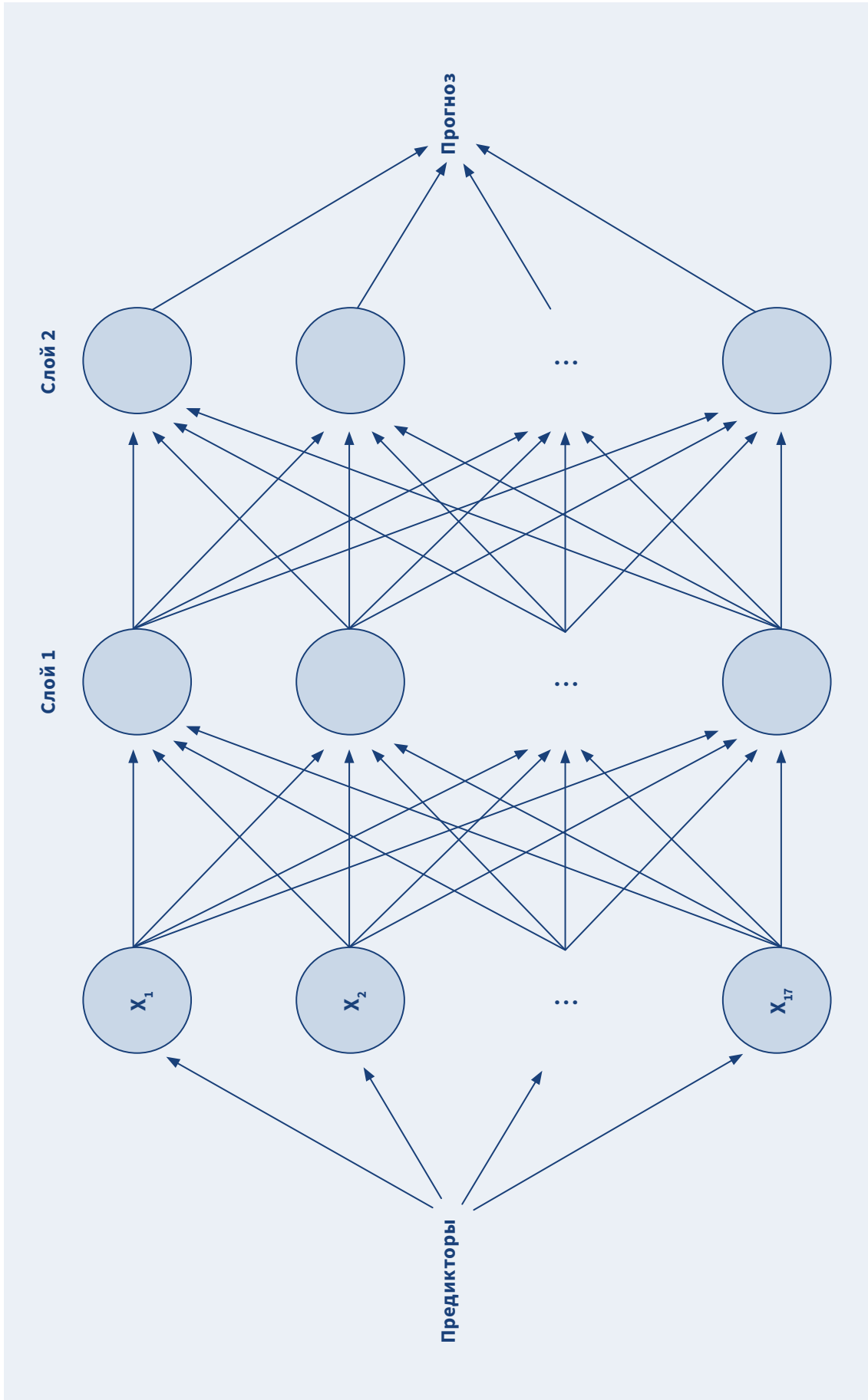
где  $X'$  — транспонированная матрица предикторов размерностью  $(17 \times N)$ ;  
 $N$  — количество наблюдений в каждом временном ряду или длина временного ряда;  
 $f_1$  и  $f_2$  — специальные функции, которые применяются для преобразования  $h_1$  и  $h_2$  соответственно.

В существующих библиотеках для Neural Network в качестве специальных функций применяются  $f_1 = 'ReLU'$ ,  $f_2 = 'linear'$ <sup>10</sup>.

5. Имея значения для  $h_1$ ,  $z_1$ ,  $h_2$  и  $\hat{y}$  осуществляется оптимизация, целью которой является обновление матрицы весов и, соответственно, пересчет  $h_1$ ,  $z_1$ ,  $h_2$  и  $\hat{y}$ . Это осуществляется с помощью алгоритма обратного распространения ошибки (backpropagation) (Chong and Zak, 2013; Hastie et al., 2009).

<sup>10</sup> Neural Networks Construction. Доступно на: <https://dafriedman97.github.io/mlbook/content/c7/construction.html>.

Рисунок 5. Визуализация нейронной сети с 17 предикторами и двумя слоями



Источник: разработка авторов.

## 2. Данные

В рамках настоящего исследования использовались временные ряды квартальных макроэкономических показателей за период с 2004Q2 по 2024Q1 (всего по 80 наблюдений). Полный список показателей представлен в [Приложении 1](#). Краткосрочному прогнозу подлежат квартальные темпы прироста реального ВВП, остальные показатели являются объясняющими (предикторами).

Перед моделированием была осуществлена предварительная обработка используемых временных рядов, которая состояла из следующих этапов:

- 1) сезонная корректировка временных рядов (при выявлении сезонности);
- 2) логарифмирование временных рядов всех показателей, кроме относительных;
- 3) приведение временных рядов к стационарному виду посредством взятия первых разностей;
- 4) аппроксимация Мариано — Мурасава ([Mariano and Murasawa, 2003](#)) для перевода ряда макропоказателей (например, инфляции, номинальной процентной ставки по вновь привлеченным депозитам в национальной валюте и др.) из месячных в квартальные.

Отметим, что для оценивания модели AR применялся только временной ряд темпов прироста реального ВВП.

Для оценивания моделей VAR и BVAR применялись три переменные: темпы прироста реального ВВП, инфляция и номинальная процентная ставка по вновь привлеченным депозитам в национальной валюте.

Для оценивания по методам и алгоритмам машинного обучения использовались все макроэкономические показатели, представленные в [Приложении 1](#). Стоит отметить, что объясняющие переменные входят в модели машинного обучения с лагом.

Для осуществления расчетов по алгоритмам машинного обучения был использован современный объектно-ориентированный язык программирования Python.

### 3. Результаты экспериментов

В рамках данного исследования для обеспечения сопоставимости рассматриваемых традиционных методов с алгоритмами машинного обучения использовались предикторы с лагом  $t-1$ .

Спецификация используемой **AR-модели** имеет следующий вид:

$$\Delta \ln rGDP_t = \beta_0 + \beta_1 \Delta \ln rGDP_{t-1} + \varepsilon_t, \quad (17)$$

где  $\beta_0, \beta_1$  — неизвестные параметры модели.

Спецификация **VAR-модели**:

$$\begin{cases} \Delta \ln rGDP_t = \alpha_{11} + \alpha_{12} \Delta \ln rGDP_{t-1} + \alpha_{13} \Delta \ln CPI_{t-1} + \alpha_{14} \Delta ddepir_{t-1} + \varepsilon_{1t}, \\ \Delta \ln CPI_t = \alpha_{21} + \alpha_{22} \Delta \ln rGDP_{t-1} + \alpha_{23} \Delta \ln CPI_{t-1} + \alpha_{24} \Delta ddepir_{t-1} + \varepsilon_{2t}, \\ \Delta ddepir_t = \alpha_{31} + \alpha_{32} \Delta \ln rGDP_{t-1} + \alpha_{33} \Delta \ln CPI_{t-1} + \alpha_{34} \Delta ddepir_{t-1} + \varepsilon_{3t}, \end{cases} \quad (18)$$

где  $\alpha_{ij}$  — неизвестные параметры модели.

При моделировании **BVAR-модели** использовались такие же показатели, как в описанной выше VAR-модели, при этом дополнительно применялась априорная информация, предложенная в [Litterman \(1980\)](#)<sup>11</sup>.

Спецификации моделей, оцениваемые при помощи **Ridge**, **LASSO** и **Elastic Net**, совпадают и имеют следующий вид:

$$\Delta \ln rGDP_t = \alpha_0 + \alpha_1 \Delta \ln ip_{t-1} + \alpha_2 \Delta \ln inv_{t-1} + \dots + \alpha_{17} \Delta fdepir_{t-1} + \varepsilon_t, \quad (19)$$

где  $\alpha_i$  — неизвестные параметры модели.

Оценивание параметров регрессии  $\alpha_i$  осуществлялось для различных значений коэффициента регуляризации  $\lambda$ , который менялся в пределах от 0,1 до 120 000. Значение параметра  $\lambda$  каждый раз пересчитывалось одновременно с обновлением временных рядов. В итоге применено то значение коэффициента  $\lambda$ , которое минимизирует ошибку прогноза.

При использовании ансамблевых методов машинного обучения (**Boosting**, **Bagging** и **Random Forest**) точность прогноза зависит от количества заданных

<sup>11</sup> Для контроля важности априорной информации применялись параметры  $w = 0,3$  (к матрице стандартных отклонений априорных параметров модели) и  $d = 1$  (к лаговым значениям модели).

бутстрэп-повторений<sup>12</sup> и выбранной глубины дерева решений<sup>13</sup>. В данном случае в результате экспериментов выявлено, что при использовании алгоритмов Boosting и Bagging минимальное значение ошибки прогноза достигается при глубине дерева решений 4 и 2 слоя соответственно. При использовании алгоритма Random Forest к указанному выше добавляется и варьирование количеством предикторов. В результате минимальная ошибка прогноза наблюдалась при использовании глубины дерева решений, равной 4 слоям, а также 4 из 17 потенциальных предикторов.

При использовании алгоритма машинного обучения **Neural networks** точность прогноза зависит от количества заданных скрытых слоев (*hidden layers*)<sup>14</sup>. Наиболее точный прогноз получен при использовании 7 скрытых слоев.

Для сравнения прогностических свойств рассматриваемых методов прогнозирования проведен регрессионный эксперимент, в частности рекурсивная схема регрессионного эксперимента, состоящая из следующих этапов:

1. Вся выборка данных первоначально делится на две части — обучающую (первые 70% наблюдений) и тестируемую (последние 30% наблюдений). В результате в обучающую выборку первоначально вошли наблюдения за период с 2004Q2 по 2018Q1, а в тестируемую — за период с 2018Q2 по 2024Q1.
2. Основываясь на фактических временных рядах обучающей выборки осуществляется оценивание модели.
3. С помощью оцененной модели производится прогноз на один квартал вперед.
4. Увеличивается длина обучающей выборки на один квартал и заново оцениваются параметры модели.
5. Используя переоцененную модель с учетом увеличенной выборки, производится прогноз на один квартал вперед.
6. Аналогично шагам 4 и 5 происходит увеличение выборки, переоценка модели и построение прогноза на один квартал вперед до тех пор, пока выборка не достигнет периода с 2004Q2 по 2023Q4.
7. В результате полученные 24 точечных прогноза темпов прироста реального ВВП сравниваются с фактическими данными и рассчитываются два индекса, которые являются функциями потерь (*loss functions*):

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}}, \quad (20)$$

$$MAD = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|, \quad (21)$$

<sup>12</sup> Согласно результатам экспериментов, 100 и более бутстрэп-повторений достаточно для получения стабильных прогнозов.

<sup>13</sup> Увеличение количества ветвей не гарантирует оптимального прогноза с точки зрения минимизации ошибки прогноза. В рамках данного исследования осуществлялся подбор глубины дерева решений в пределах от 2 до 10 слоев.

<sup>14</sup> В рамках данного исследования осуществлялся подбор количества скрытых слоев от 2 до 10.

где  $y_t$  и  $\hat{y}_t$  — фактические и оцененные по различным моделям значения  $rGDP_t$ ;

$n$  — общее количество оценок, в данном случае  $n = 24$ .

С учетом результатов расчетов  $RMSE_i$  и  $MAD_i$  (Таблица 1) получены следующие выводы:

- AR-модель по качеству прогноза превосходит VAR- и BVAR-модели;
- все рассматриваемые методы и алгоритмы машинного обучения (за исключением Boosting, согласно  $MAD$ , и Bagging, согласно  $RMSE$  и  $MAD$ ) по качеству прогноза превосходят AR-модель;
- по результатам статистики, предложенной Diebold and Mariano (1995), различия между прогнозами AR-модели и иных анализируемых методов (за исключением VAR-модели) не существенны.

**Таблица 1. Результаты оценки точности краткосрочного прогноза темпов прироста реального ВВП с помощью сравниваемых моделей**

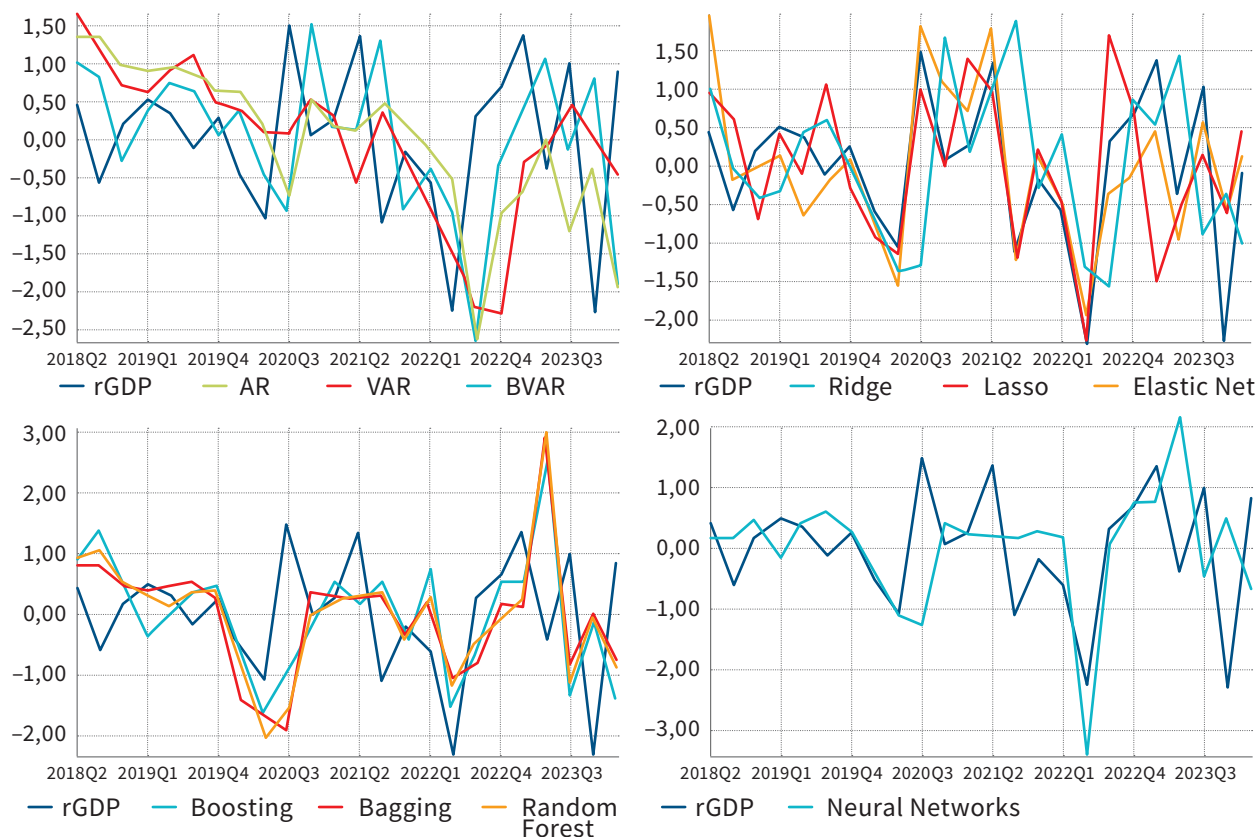
Модели	RMSE	RMSE (по отнош. к AR)	DM-stat.	MAD	MAD (по отнош. к AR)	DM-stat.
AR	2,121	–	–	1,601	–	–
VAR	2,177	1,026	1,304	1,703	1,064	2,184*
BVAR	2,183	1,029	0,736	1,698	1,061	1,160
Ridge	1,893	0,893	–1,325	1,485	0,928	–0,732
LASSO	1,832	0,864	–1,613	1,414	0,883	–1,209
Elastic Net	1,840	0,868	–1,639	1,424	0,889	–1,114
Boosting	2,113	0,996	–0,040	1,649	1,030	0,267
Bagging	2,127	1,003	0,026	1,623	1,014	0,132
Random Forest	2,090	0,985	–0,185	1,596	0,997	–0,034
Neural Network	1,914	0,902	–0,850	1,469	0,918	–0,852

\* Существенна на 5%-ном уровне значимости.

При сравнении результатов прогноза темпов прироста реального ВВП, полученных анализируемыми методами, выявлено следующее (Рисунок 6):

- при использовании моделей машинного обучения не наблюдается лагового сдвига, характерного для AR-, VAR- и BVAR-моделей;
- пики и спады темпов прироста реального ВВП наиболее точно прогнозируются с помощью методов регуляризации (Ridge, LASSO, Elastic Net) и Neural Networks. Это может объясняться тем, что традиционные методы ориентированы на лаговые значения, а прогнозы по моделям машинного обучения — на всю обучающую выборку.

Рисунок 6. Сравнение результатов прогноза темпов прироста реального ВВП



Источник: расчеты авторов.

Отвечая на вопрос «какой из алгоритмов использовать в практических целях?», следует отметить, что это должен быть тот алгоритм машинного обучения, который обеспечивает минимальное значение  $RMSE$  и  $MAD$ . В данном случае, согласно Таблице 1, наиболее точным является алгоритм LASSO. Однако со временем более точным может стать любой из рассматриваемых методов. В этой связи можно применять комбинированный прогноз. Комбинированный прогноз (Richardson et al., 2018) можно осуществлять путем взвешивания на основе:

- равных весов;
- весов, полученных с помощью МНК;
- обратных значений функции потерь;
- обратных значений рангов функции потерь.

Согласно полученным результатам взвешивания, описанными выше методами (Таблица 2), выявлено, что все они обеспечивают более низкие значения  $RMSE$  и  $MAD$  относительно AR-модели. При этом самая высокая точность наблюдается при взвешивании с помощью МНК весов<sup>15</sup>.

<sup>15</sup> В данном исследовании при определении МНК весов осуществлялось построение регрессии без константы и ограничений на параметры в соответствии с Richardson et al. (2018).

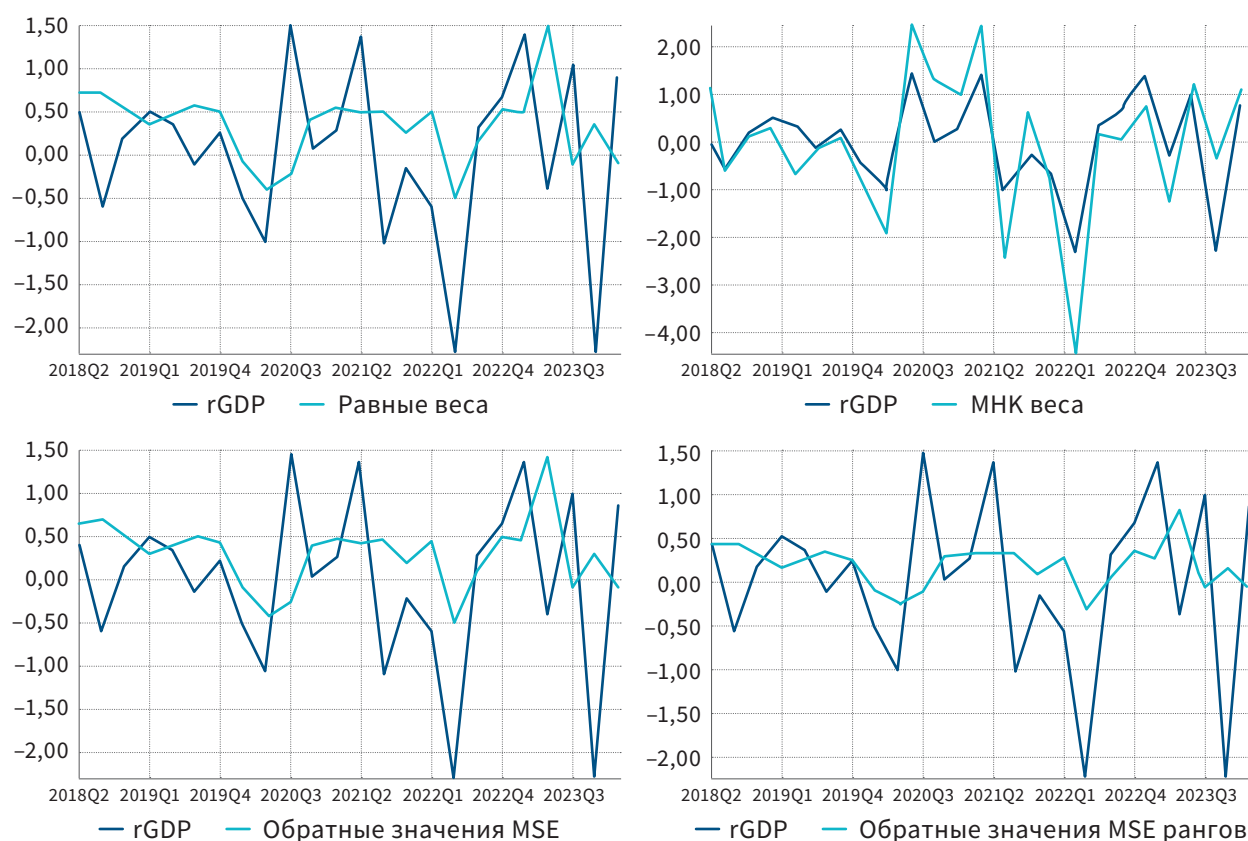
**Таблица 2. Комбинирование прогнозов темпа прироста реального ВВП**

Метод взвешивания	RMSE	RMSE (по отнош. к AR)	MAD	MAD (по отнош. к AR)
Равные веса	1,922	0,906	1,465	0,915
МНК веса	1,127	0,531*	0,763	0,477*
Обратные значения MSE	1,915	0,903	1,461	0,913
Обратные значения MSE рангов	1,887	0,890	1,457	0,910

\* Существенна на 5%-ном уровне значимости.

Согласно [Рисунку 7](#), на котором сравниваются фактические темпы прироста реального ВВП и комбинированных прогнозов, наиболее точное улавливание спадов и пиков темпов прироста реального ВВП наблюдается и в случае комбинированного прогноза по МНК весам.

**Рисунок 7. Комбинированные прогнозы темпов прироста реального ВВП**



**Примечание:** расчеты авторов.

Выводы, которые были сделаны:

1. Алгоритмы машинного обучения позволяют получить более точные прогнозы по сравнению с традиционными подходами.

2. Разница прогнозов при сравнении с AR-моделью статистически незначима, согласно статистике Диболда — Мариано;
3. Как наилучший прогноз можно использовать результаты того алгоритма, который минимизирует функцию потерь. В дальнейшем, по мере поступления новых данных, выбор модели может измениться в зависимости от тенденций в показателях  $RMSE_t$  и  $MAD_t$ .
4. В качестве прогноза можно использовать комбинированный прогноз на основе МНК весов. Причем этот подход, как показывают результаты расчетов, позволяет получить прогнозы, которые статистически значимо отличаются от прогнозов по AR-модели.

## Заключение

На текущий момент основными инструментами для краткосрочного прогнозирования макроэкономических показателей остаются традиционные эконометрические методы и модели. Однако в связи с интенсивным развитием технологий и вычислительной техники все чаще возникает спрос на применение в экономическом анализе и прогнозировании методов и алгоритмов машинного обучения.

В данном исследовании проведены эксперименты с целью 1) определения потенциала методов и алгоритмов машинного обучения повысить точность прогнозов и 2) выяснения, могут ли эти современные подходы стать как дополнением к традиционным инструментам краткосрочного прогнозирования, так и их альтернативой.

Для сравнения использовались три традиционные эконометрические модели (AR, VAR и BVAR), а также семь методов и алгоритмов машинного обучения (Ridge- и LASSO-регрессии, Elastic Net, Boosting, Bagging, Random Forest и Neural Networks).

По результатам проведенных экспериментов сделаны следующие выводы:

1. Практически все рассматриваемые в работе методы и алгоритмы машинного обучения по качеству прогноза превосходят традиционные инструменты краткосрочного прогнозирования (в части минимизации функции потерь).
2. По результатам статистики Диболда — Мариано различия между прогнозами анализируемых методов статистически не существенны.
3. При использовании методов и алгоритмов машинного обучения не наблюдается лагового сдвига, характерного для AR-, VAR- и BVAR-моделей.
4. Результаты прогнозов существенно улучшаются за счет применения комбинированного прогноза. Наибольшая точность прогнозов достигается при использовании комбинированного прогноза на основе МНК весов. В таком случае полученные результаты статистически значимо отличаются от прогнозов по AR-модели.

Таким образом, на основе проведенного исследования можно сделать вывод, что методы и алгоритмы машинного обучения целесообразно применять в качестве альтернативы или дополнения к традиционным эконометрическим методам для повышения точности краткосрочного прогноза макроэкономических показателей.

## Список литературы

Жемков, М. (2021). Краткосрочная оценка ВВП России методом комбинирования прогнозов. Банк России, Серия докладов об экономических исследованиях. Доступно на: <https://www.cbr.ru/StaticHtml/File/118606/wp-feb21.pdf>.

Alessi, L., Savona, R. (2021). Machine Learning for Financial Stability. In: Consoli, S., Reforgiato Recupero, D., Saisana, M. (eds) Data Science for Economics and Finance. Springer, Cham. Available at: [https://doi.org/10.1007/978-3-030-66891-4\\_4](https://doi.org/10.1007/978-3-030-66891-4_4).

Barbaglia, L., Consoli, S., Manzan, S., Reforgiato Recupero, D., Saisana, M., Tiozzo Pezzoli, L. (2021). Data Science Technologies in Economics and Finance: A Gentle Walk-In. In: Consoli, S., Reforgiato Recupero, D., Saisana, M. (eds) Data Science for Economics and Finance. Springer, Cham. Available at: [https://doi.org/10.1007/978-3-030-66891-4\\_1](https://doi.org/10.1007/978-3-030-66891-4_1).

Biau, O. and D'Elia, A. (2010). Euro Area GDP Forecasting Using Large Survey Datasets: A Random Forest Approach, Euroindicators working papers.

Breiman, L. and Ihaka, R. (1984). Nonlinear discriminant analysis via scaling and ACE. Technical Report, University of California Berkeley.

Breiman, L. (1996). Bagging predictors, Machine Learning 26: 123–140.

Breiman, L. (2001). Statistical Modeling: The Two Cultures, Statistical Science, Vol. 16, No. 3, pp. 199–231.

Buckmann M., Joseph A., and Robertson, H. (2021). Opening the Black Box: Machine Learning Interpretability and Inference Tools with an Application to Economic Forecasting, in S. Consoli et al. (eds.), Data Science for Economics and Finance. Available at: [https://doi.org/10.1007/978-3-030-66891-4\\_3](https://doi.org/10.1007/978-3-030-66891-4_3).

Chakraborty, C. and Joseph, A. (2017). Machine learning at central banks, Bank of England Staff Working Paper No. 674.

Chong, E. and Zak, S. (2013). An introduction to optimization. A John Wiley and Sons, Fourth Edition.

Chuku, C., Simpasa, A., and Oduor, J. (2019). Intelligent forecasting of economic growth for developing economies, International Economics, vol. 159, pp. 74–93. Available at: <https://doi.org/10.1016/j.inteco.2019.06.001>.

Desai, A. (2023). Machine learning for economics research: when, what and how. Bank of Canada, Staff Analytical Note 2023-16. Available at: <https://www.bankofcanada.ca/2023/10/staff-analytical-note-2023-16/>.

Diebold, F. X. and Mariano, R. S. (1995). Comparing Predictive Accuracy. Journal of Business & Economic Statistics 13(3), 253–263.

Efron, B. and Tibshirani, R. (1993). An introduction to the Bootstrap. Chapman and Hall, London.

Freund, Y. and Shapire, R. (1997). A decision-theoretic generalization of online learning and an application to boosting. *Journal of Computer and System Sciences* 55: 119–139.

Gupta, R. and Kabundi, A. (2009). A large factor model for forecasting macroeconomic variables in South Africa. Working paper 137, University of Cape Town.

Hamilton, J. (1994). *Time series analysis*. Princeton University Press.

Hastie, T., Tibshirani, R. & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Stanford, CA: Stanford University.

Hoerl, A. E. and Kennard, R. (1970). Ridge regression: biased estimation for nonorthogonal problems, *Technometrics* 12: 55–67.

Litterman, R. (1980). A Bayesian procedure for forecasting with vector autoregressions. Manuscript, MIT, Department of Economics.

Mariano, R. and Murasawa, Y. (2003). A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series. *Journal of Applied Econometrics*, vol. 18, issue 4, pp. 427–443.

Pirschel, I. and Wolters, M. (2014). Forecasting German key macroeconomic variables using large dataset methods, *Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik — Session: Forecasting*, No. B16-V4, ZBW — Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationzentrum Wirtschaft, Kiel und Hamburg.

Poghosyan, K. (2015). A Comparison of Different Short Term Macroeconomic Forecasting Models: Evidence from Armenia, *Journal of Central Banking Theory and Practice*, 2016, 2, pp. 81–99. DOI: 10.1515/jcbtp-2016-0012.

Poghosyan, K. and Magnus, J. R. (2012). WALs Estimation and Forecasting in Factor-based Dynamic Models with an Application to Armenia, *International Econometric Review*, Econometric Research Association, vol. 4(1), pp. 40–58. Available at: <https://ideas.repec.org/a/erh/journal/v4y2012i1p40-58.html>.

Richardson, A., Florenstein Mulder, T., Vehbi, T. (2018). Nowcasting New Zealand GDP using machine learning algorithms. 2018-47, Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, the Australian National University. Available at: <https://ideas.repec.org/s/een/camaaa.html>.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society, Series B* 58: 267–288.

Tiffin, A. (2016). *Seeing in the Dark: A Machine-Learning Approach to Nowcasting in Lebanon*, IMF Working Paper.

Tkacz, G. and Hu, S. (1999). Forecasting GDP Growth Using Artificial Neural Networks, Bank of Canada Working Paper No. 99–3.

Vaish R., Dwivedi, U. D., Tewari, S., Tripathi, S. M. (2021). Machine learning applications in power system fault diagnosis: Research advancements and perspectives, *Engineering Applications of Artificial Intelligence*, vol. 106. Available at: <https://doi.org/10.1016/j.engappai.2021.104504>.

## Приложение

Таблица. Список использованных временных рядов

Наименование показателя	Краткое обозначение	Источник	Сезонная корректировка	Трансформация
<b>Прогнозируемый показатель</b>				
ВВП в среднегодовых ценах 2018 года, млн рублей	$rGDP_t$	Белстат	Да	$\ln, \Delta$
<b>Объясняющие показатели (предикторы)</b>				
Объем промышленного производства в среднегодовых ценах 2018 года, млн рублей	$ip_t$	Белстат	Да	$\ln, \Delta$
Объем инвестиций в основной капитал в среднегодовых ценах 2018 года, млн рублей	$inv_t$	Белстат	Да	$\ln, \Delta$
Объем розничного товарооборота в среднегодовых ценах 2018 года, млн рублей	$rt_t$	Белстат	Да	$\ln, \Delta$
Индекс потребительских цен (100 = 2004Q1)	$CPI_t$	Белстат	Да	$\ln, \Delta$
Индекс цен производителей промышленной продукции (100 = 2004Q1)	$PPI_t$	Белстат	Нет	$\ln, \Delta$
Ежемесячная цена российского природного газа, долларов за миллион британских тепловых единиц	$ngas_t$	IndexMundi	Нет	$\ln, \Delta$
Цена на нефть марки Brent, долларов за баррель	$oilpr_t$	IndexMundi	Нет	$\ln, \Delta$
Экспорт товаров и услуг в текущих ценах, млн долларов	$exp_t$	Белстат	Да	$\ln, \Delta$
Импорт товаров и услуг в текущих ценах, млн долларов	$ipm_t$	Белстат	Да	$\ln, \Delta$
Денежная база (среднее за месяц), млн рублей	$MO_t$	НБРБ	Да	$\ln, \Delta$
Срочные депозиты в нац. валюте (среднее за месяц), млн рублей	$ddep_t$	НБРБ	Да	$\ln, \Delta$
Срочные депозиты в ин. валюте (среднее за месяц), млн рублей	$fdep_t$	НБРБ	Да	$\ln, \Delta$
Широкая денежная масса (среднее за месяц), млн рублей	$BM_t$	НБРБ	Да	$\ln, \Delta$
Кредитные вложения банков в частную собственность, млн рублей	$loanpr_t$	НБРБ	Да	$\ln, \Delta$
Кредитные вложения банков в государственную собственность, млн рублей	$loangov_t$	НБРБ	Да	$\ln, \Delta$
Средняя номинальная процентная ставка по вновь привлеченным депозитам в нац. валюте, %	$ddepir_t$	НБРБ	Нет	$\Delta$
Средняя номинальная процентная ставка по вновь выданным кредитам (без МБК) в нац. валюте, %	$fdepir_t$	НБРБ	Нет	$\Delta$

**Примечание:**  $\ln$  — логарифмирование показателя,  $\Delta$  — взятие первых разностей.



**Совместный Рабочий документ ЕФСР и Института ЦАРЭС (RU/EN)**

**Анализ процентного риска стран для долговой и бюджетной устойчивости: потенциал использования обязательств с плавающими расходами**

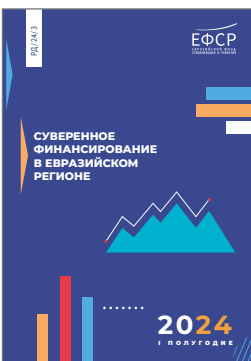
В документе рассмотрен процентный риск, связанный с недавним ростом ставок USD LIBOR (SOFR) и EURIBOR, а также потенциал использования инструментов с плавающими расходами на обслуживание и погашение во внутренних долговых портфелях.



**Рабочий документ РД/24/4 (RU/EN)**

**Система раннего предупреждения ЕФСР: методология разработки инструментов для идентификации валютных кризисов**

В документе представлены методология и пошаговый алгоритм разработки инструментов для идентификации дисбалансов (кризисов) и стрессовых ситуаций в экономике. Основной акцент сделан на своевременном выявлении нарастания напряженности на валютном рынке.



**Рабочий документ РД/24/3 (RU/EN)**

**Суверенное финансирование в Евразийском регионе: I полугодие 2024 года**

Документ посвящен мониторингу суверенного финансирования в Евразийском регионе за I полугодие 2024 года.



**Рабочий документ РД/24/2 (RU/EN)**

**Суверенное финансирование в Евразийском регионе: тенденции и направления**

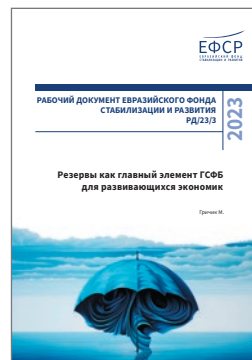
Документ опирается на базу данных, которую ведет ЕФСР, и посвящен мониторингу суверенного финансирования в Евразийском регионе за 2023 год.



**Рабочий документ РД/24/1 (RU/EN)**

**Оценка потенциального влияния стихийных бедствий на долговую устойчивость Армении, Кыргызстана и Таджикистана**

В документе представлен алгоритм действий, который может быть использован для оценки влияния стихийных бедствий на макроэкономические параметры и долговую устойчивость государств.



**Рабочий документ РД/23/3 (RU/EN)**

**Резервы как главный элемент ГСФБ для развивающихся экономик**

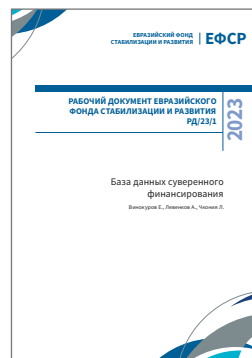
В документе оценены факторы, влияющие на принятие решения развивающимися экономиками об источнике антикризисной поддержки. Анализ показал, что международные резервы являются наиболее востребованным инструментом из всех элементов ГСФБ.



**Рабочий документ РД/23/2 (RU/EN)**

**Суверенное финансирование в Евразийском регионе: водный сектор и гидроэнергетика**

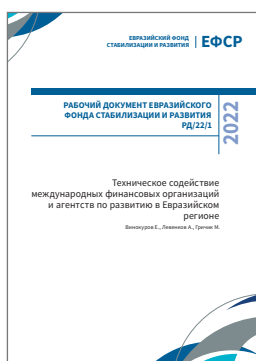
В данном Рабочем документе ЕФСР проведен количественный и качественный анализ операций суверенного финансирования в водно-энергетическом секторе (ВЭС) в 11 странах региона с 2008 по первое полугодие 2023 года.



**Рабочий документ РД/23/1 (RU/EN)**

**База данных суверенного финансирования**

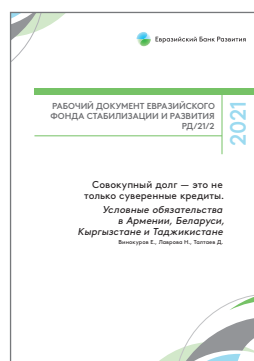
В данном Рабочем документе ЕФСР представлена методология построения Базы данных суверенного финансирования (SFD), а также проведен количественный и качественный анализ операций суверенного финансирования в 11 странах региона с 2008 по 2022 год.



**Рабочий документ РД/22/1 (RU/EN)**

**Техническое содействие международных финансовых организаций и агентств по развитию в Евразийском регионе**

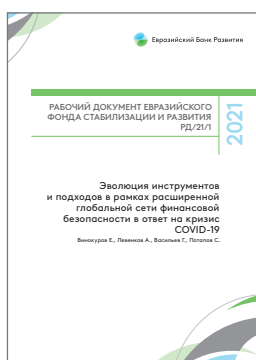
В исследовании проведен количественный и качественный анализ проектов технического содействия в 11 странах региона в разрезе отдельных тематических направлений и институтов.



**Рабочий документ РД/21/2 (RU/EN)**

**Совокупный долг — это не только суверенные кредиты. Условные обязательства в Армении, Беларуси, Кыргызстане и Таджикистане**

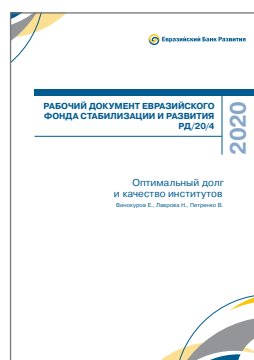
Исследование отвечает на вопрос о потенциальных рисках реализации прямых и косвенных условных обязательств и их влиянии на бюджетные и долговые позиции 4 стран — получателей средств ЕФСР. Особое внимание уделяется государственным предприятиям и их роли в долговых позициях стран.



**Рабочий документ РД/21/1 (RU/EN)**

**Эволюция инструментов и подходов в рамках расширенной Глобальной сети финансовой безопасности в ответ на кризис COVID-19**

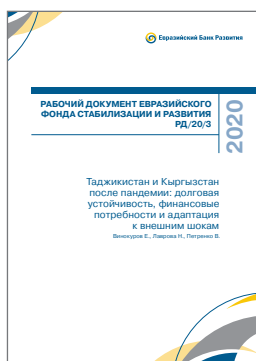
Авторы проанализировали реакцию элементов ГСФБ на кризис, вызванный пандемией COVID-19, в масштабе мировой экономики и на уровне региона операций ЕФСР.



**Рабочий документ РД/20/4 (RU/EN)**

**Оптимальный долг и качество институтов**

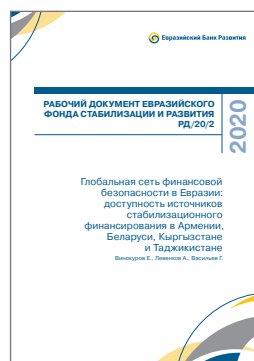
Текущий кризис, вызванный пандемией COVID-19, поставил многие страны перед выбором – стимулировать экономическое развитие за счет наращивания долговых обязательств или сосредоточиться на бюджетной устойчивости в условиях макроэкономической неопределенности.



**Рабочий документ РД/20/3 (RU/EN)**

**Таджикистан и Кыргызстан после пандемии: долговая устойчивость, финансовые потребности и адаптация к внешним шокам**

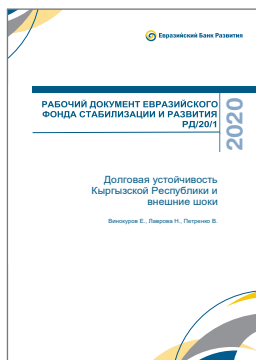
Пандемия COVID-19 продемонстрировала чувствительность экономики и долговых позиций стран к целому ряду потрясений.



**Рабочий документ РД/20/2 (RU/EN)**

**Глобальная сеть финансовой безопасности в Евразии: доступность источников стабилизационного финансирования в Армении, Беларуси, Кыргызстане и Таджикистане**

В аналитическом документе учтены шесть источников антикризисного финансирования.



**Рабочий документ РД/20/1 (RU/EN)**

**Долговая устойчивость Кыргызской Республики и внешние шоки**

Документ содержит методологический и эмпирический анализ долговой устойчивости Кыргызстана. В работе дана оценка эффекта различных видов шоков на долговую устойчивость страны.



**Рабочий документ РД/19/2 (RU/EN)**

**Достижение целей стабилизации и развития в рамках одного мандата: опыт Евразийского фонда стабилизации и развития**

В этом исследовании проанализирован опыт ЕФСР, свидетельствующий о том, что в контексте стран с низким уровнем дохода мандат РФМ в области стабилизации может быть дополнен задачами по развитию.



Алгоритмы машинного обучения для краткосрочного  
прогнозирования темпов прироста реального ВВП

Цукарев Т., Погосян К., Лемба К.

**Евразийский фонд стабилизации и развития (ЕФСР)** объемом более 9 млрд долларов США учрежден 9 июня 2009 года правительствами Армении, Беларуси, Казахстана, Кыргызской Республики, России и Таджикистана. Целями ЕФСР является содействие государствам-участникам в преодолении последствий глобального финансового кризиса, в обеспечении их экономической и финансовой стабильности и поддержке интеграционных процессов в регионе. Детальная информация о ЕФСР доступна по ссылке: [efsd.org](http://efsd.org).

**Серия рабочих документов ЕФСР** является основным форматом публичной аналитики Фонда. Рабочие документы ЕФСР представляют собой аналитические исследования Фонда в области глобальных, региональных и страновых экономических тенденций, экономического моделирования, макроэкономического анализа, отраслевого анализа, мировой финансовой архитектуры и иных вопросов. Публикации ЕФСР доступны по ссылке: [efsd.org/research](http://efsd.org/research).

**Контакты**

Москва

Чистопрудный бульвар, 17, стр. 1  
101000, Российская Федерация

Тел.: +7 (495) 645 04 45

Факс: +7 (495) 645 04 41

Веб: [efsd.org](http://efsd.org)



[www.efsd.org](http://www.efsd.org)