

Наукастинг ВВП: от традиционных эконометрических моделей к машинному обучению

Цукарев Т., Погосян К., Лемба К., Гришин Д.

РД/25/1

Ключевые слова: наукастинг, темп прироста реального ВВП, дефлятор ВВП, Bridge-уравнение, MIDAS, факторная модель, машинное обучение.

Коды JEL: C22, C52, C53, C55.

Контакты авторов:

Тарас Цукарев, Департамент по аналитической работе ЕФСР: ttsukarev@efsd.org

Карен Погосян, Департамент по аналитической работе ЕФСР: kpoghosyan@efsd.org

Кирилл Лемба, Департамент по аналитической работе ЕФСР: KLemba@efsd.org

Данил Гришин, Департамент по аналитической работе ЕФСР: dgrishin@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 Совета директоров ЕФСР, Совета управляющих ЕФСР или государств — участников 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 литературы | 6 |
| 2. Обзор моделей | 10 |
| 3. Данные и программное обеспечение | 20 |
| 4. Результаты экспериментов | 21 |
| Заключение | 29 |
| Список литературы | 30 |
| Приложение 1 | 33 |
| Приложение 2 | 34 |

Список сокращений

| | |
|----------------|---|
| Армстат | Статистический комитет Республики Армения |
| Белстат | Национальный статистический комитет Республики Беларусь |
| ВВП | валовой внутренний продукт |
| доллар | доллар США |
| ЕФСР | Евразийский фонд стабилизации и развития |
| МНК | метод наименьших квадратов |
| НБРБ | Национальный банк Республики Беларусь |
| ЦБ РА | Центральный банк Республики Армения |
| AIC | Akaike's information criterion |
| ARDL | autoregressive distributed lag |
| ARIMA | autoregressive integrated moving average |
| DFM | dynamic factor model |
| DM | Diebold — Mariano |
| GRU | gated recurrent unit |
| MAD | mean absolute deviation |
| MIDAS | mixed data sampling |
| MF | mixed-frequency |
| MLP | multilayer perceptron |
| MSE | mean squared error |
| LASSO | least absolute shrinkage and selection operator |
| LSTM | long short-term memory |
| RMSE | root mean squared error |
| RNN | recurrent neural network |
| RSS | residuals sum of squares |
| SVM | support vector machine |
| SVR | support vector regression |
| VAR | vector autoregression |

Благодарность

Авторы выражают благодарность Гергели Акошу Ганичу (экономисту-исследователю Управления по вопросам финансовой стабильности, регулирования и санации банков, Банк Испании), Дине Хенг (старшему экономисту Департамента Ближнего Востока и Центральной Азии, Международный валютный фонд), а также сотрудникам ЕФСР за ценные комментарии и рекомендации. Ответственность за любые ошибки и недочеты остается на авторах.

Введение

Оценка текущей ситуации в экономике, как правило, осложняется тем, что значительная часть ключевых статистических данных публикуется с существенной задержкой и подлежит последующим пересмотрам (например, ВВП и его структура, показатели государственного бюджета, платежного баланса и т. д.). В отличие от метеорологов, которые опираются на актуальные данные о погоде и строят краткосрочные прогнозы, макроэкономисты вынуждены анализировать не только текущее состояние экономики, но и оценить динамику основных показателей как минимум за предыдущие один-два месяца. Для преодоления этой проблемы применяются методы «наукастинга» (от англ. *nowcasting*), позволяющие оценивать макроэкономические показатели в режиме, приближенном к реальному времени.

Под наукастингом понимается оценка самого недавнего прошлого, настоящего и прогноз самого близкого будущего (BańBura, 2010; Малюгин, 2024). Наукастинг применяется в отношении ключевых макроэкономических переменных, которые представляются с низкой частотой (например, на ежеквартальной основе) и публикуются со значительным временным лагом. Для получения «ранних оценок» таких макроэкономических показателей эксперты в области прогнозирования используют информацию с более высокой частотностью и более оперативной публикацией. Например, реальный ВВП, являющийся основным индикатором состояния экономики, доступен в большинстве случаев лишь с квартальной периодичностью, и первая оценка публикуется с лагом около двух-трех месяцев. В то же время существует ряд переменных, тесно связанных с ВВП, которые формируются с гораздо более высокой частотой (еженедельно или ежемесячно) и публикуются с меньшей задержкой: например, данные по промышленному производству, показатели финансового сектора или результаты различных опросов. Как раз такие данные могут быть использованы для формирования ранних оценок ВВП при помощи методов наукастинга. Ключевой аспект процесса — использование актуальной высокочастотной информации в условиях несинхронного раскрытия данных, характеризующихся различными задержками, что приводит к формированию наборов наблюдений с так называемым «неровным» краем.

Анализ международной практики демонстрирует, что в арсенале у исследователей и экспертов, занимающихся наукастингом макроэкономических показателей, широкий перечень инструментов, которые, как правило, основаны на эконометрическом анализе временных рядов. Наибольшую популярность, особенно в центральных банках, получили такие методы и модели, как Bridge-уравнения, модель MIDAS, а также динамические факторные модели (DFM) (Forni and Marcellino, 2013; Chernis and Sekkel, 2017). Постоянное развитие этих моделей, обусловленное методологическими инновациями и интеграцией различных источников данных, способствует повышению точности и надежности экономических оценок в режиме реального времени.

В последние годы все большую популярность приобретают методы (алгоритмы) машинного обучения (Ridge, LASSO, Elastic Net, Boosting, Bagging, Random Forest, Support Vector Machine, Multilayer Perceptron, Recurrent Neural Network и др.), которые выступили своеобразной альтернативой традиционным эконометрическим моделям (Desai, 2023). При этом алгоритмы машинного обучения применяются как для решения задачи регрессии¹, так и для задачи классификации.

¹ В данной работе акцентируется внимание на применении этих алгоритмов для целей регрессии в области наукастинга (Medeiros et al., 2021; Richardson et al., 2021).

Цель настоящего исследования — оценить потенциал методов и алгоритмов машинного обучения в повышении точности оценок макроэкономических показателей и определить, могут ли они стать эффективным дополнением к традиционным инструментам наукастинга или даже их альтернативой.

В данной работе рассматриваются три традиционные модели для наукастинга — Bridge-уравнения, модель MIDAS и факторная модель. В качестве сравнительной альтернативы к ним взяты такие методы и алгоритмы машинного обучения, как Ridge- и LASSO-регрессии, Elastic Net, Boosting, Bagging, Random Forest, SVM, MLP и RNN.

Для наукастинга на основе традиционных моделей и алгоритмов машинного обучения использованы фактические макроэкономические временные ряды по Республике Армения и Республике Беларусь за период с 2002 по 2024 год. В исследовании применялись ряды с месячной частотностью: для Армении — 22 показателя, для Беларуси — 20. В качестве зависимых переменных выступили темп прироста реального ВВП и дефлятор ВВП, имеющие квартальную частотность. Большая часть временных рядов была подвергнута первоначальной статистической обработке, а именно сезонному сглаживанию, логарифмированию и взятию первых разностей для устранения нестационарности.

С целью определения наиболее эффективного с точки зрения прогностических свойств метода наукастинга был проведен ряд экспериментов. Их суть в следующем: общее количество наблюдений во временных рядах делится на две части — обучающую и тестовую выборки. Первая используется для оценки параметров моделей, вторая — для сравнения точности оценок (прогнозов), полученных с помощью традиционных моделей и методов машинного обучения, с фактическими значениями. Генерация прогнозов осуществляется по рекурсивной схеме регрессионного эксперимента (Poghosyan and Magnus, 2012). Наилучшим считается тот инструмент, который минимизирует выбранную для этих целей функцию потерь.

Результаты экспериментов показывают, что оценки, полученные с помощью алгоритмов машинного обучения, по точности превосходят прогнозные оценки традиционных моделей. Наиболее эффективными инструментами оказались LASSO-регрессия, Boosting, Random Forest, Random Forest, SVM и RNN.

Более того, статистический тест Диболда — Мариано подтверждает, что прогнозные оценки, полученные с помощью LASSO-регрессии, SVM и RNN, существенно отличаются от результатов традиционной модели Bridge-уравнения. При этом расчеты продемонстрировали, что комбинирование различных прогнозов позволяет дополнительно повысить точность наукастинга по сравнению с использованием отдельных методов.

Данное исследование имеет следующую структуру. В первом разделе представлен краткий обзор литературы, посвященной исследуемой проблеме. Во втором разделе приведено описание использованных в работе моделей и алгоритмов. Третий раздел содержит детальный анализ исходных данных, включая методы предварительной статистической обработки временных рядов. В четвертом разделе описываются эксперименты и полученные на их основе результаты. В заключении сформулированы основные выводы и практические рекомендации.

1. Обзор литературы

В данном разделе рассматриваются две группы исследовательских публикаций, посвященных наукастингу макроэкономических переменных. Первая группа концентрирует внимание на традиционных моделях, основанных на эконометрическом анализе временных рядов (Bridge-уравнения, модель MIDAS, динамические факторные модели). Вторая группа охватывает современные альтернативные инструменты, основанные на методах и алгоритмах машинного обучения. В последние годы именно они демонстрируют растущую популярность за счет способности выявлять сложные нелинейные зависимости в данных и работать с большими массивами информации.

Одним из наиболее простых и широко применяемых инструментов наукастинга являются Bridge-уравнения, которые позволяют оценивать текущие значения низкочастотных макроэкономических показателей (например, квартального ВВП) на основе высокочастотных индикаторов (таких как месячные данные промышленного производства, розничных продаж или индексов деловой активности). Благодаря своей простоте и интерпретируемости, Bridge-уравнения активно используются центральными банками и другими институтами, занимающимися экономическим прогнозированием, в качестве ключевого элемента наукастинговых систем для мониторинга текущей динамики ВВП и других важных макропеременных. Например, [Baffigi et al. \(2004\)](#) использовали Bridge-уравнения для наукастинга ВВП Германии, Франции, Италии и еврозоны. Авторы демонстрируют, что применение национальных моделей и последующая их агрегация позволяют лучше учитывать специфические особенности отдельных стран, это способствует повышению точности прогнозов для всей еврозоны. Таким образом, статья подчеркивает эффективность Bridge-уравнений в контексте наукастинга и их преимущество перед другими базовыми моделями краткосрочного прогнозирования ВВП, такими как ARIMA, VAR или структурные модели.

Методология наукастинга нашла развитие в работе [Clements and Galvao \(2008\)](#). Авторы предложили интегрировать данные разной частотности в прогнозные модели. Они рассматривают построение различных Bridge-уравнений, которые связывают высокочастотные индикаторы с целевой переменной — ростом ВВП США. [Clements and Galvao \(2008\)](#) подкрепляют эффективность своего подхода эмпирическими доказательствами, проводя прогнозные эксперименты с использованием данных в режиме реального времени. Сравнивая точность прогнозов моделей, включающих данные смешанной частоты, с моделями, базирующимися исключительно на низкочастотных данных, авторы демонстрируют возможное улучшение точности прогнозов за счет включения информации в режиме реального времени. Подчеркивая преимущества использования более широкого набора источников данных, они отмечают важность адаптации прогнозных методик к условиям цифровой эпохи, характеризующейся все более доступными высокочастотными данными.

Несмотря на широкое распространение, использование Bridge-уравнений сопряжено с рядом трудностей ([Жемков, 2021](#)). Во-первых, для получения оценки (прогноза) искомой низкочастотной переменной требуется прогноз всех экзогенных высокочастотных переменных, включенных в модель. Эти прогнозы могут основываться как на простых моделях (например, авторегрессионных), так и на более сложных. Однако в любом случае существует риск накопления ошибки, связанной с неточностью прогноза экзогенных факторов. Во-вторых, в практике затруднен выбор наиболее информативных переменных: определение ключевых предикторов среди множества доступных индикаторов

представляет собой нетривиальную задачу. В связи с этим в современной литературе наметилась тенденция к использованию так называемого «связывания через факторы» (*bridging with factors*), когда вместо отдельных переменных в уравнения включаются агрегированные латентные факторы, извлеченные из массива данных (Bernanke et al., 2005; Giannone et al., 2008). Такой подход стал доминирующим и считается стандартом в большинстве центральных банков и международных организаций.

Статья Giannone et al. (2008) стала одной из ключевых в становлении наукастинга как самостоятельного направления в прикладной макроэкономике. Авторы предлагают динамическую факторную модель (DFM), способную обрабатывать большие массивы данных различной частоты в условиях реального времени. Основной акцент сделан на том, как поступление новых макроэкономических индикаторов в течение квартала обновляет оценку текущего роста ВВП США. Используя данные в реальном времени, авторы демонстрируют, что модель эффективно отслеживает макроэкономическую динамику и обеспечивает значительное повышение точности прогнозов по сравнению с традиционными подходами. Работа также подчеркивает важность последовательности публикации статистических данных и показывает, что наибольший прирост точности прогнозной оценки происходит на ранних этапах квартала, благодаря публикации «мягких» (soft) индикаторов (опросов и ожиданий). Следуя Giannone et al. (2008), было опубликовано достаточное количество исследований по наукастингу макропоказателей для экономик таких стран, как Франция (Barhoumi et al., 2010), Новая Зеландия (Matheson, 2010), Норвегия (Aastveit et al., 2011), Китай (Giannone et al., 2013), Канада (Chernis and Sekkel, 2017) и др.

Например, в статье Chernis and Sekkel (2017) авторы разрабатывают и оценивают DFM для наукастинга роста ВВП Канады. Модель объединяет «жесткие» и «мягкие» индикаторы, включая значительную долю международных данных, в частности по США. В рамках экспериментов в псевдореальном времени авторы демонстрируют, что DFM превосходит по точности прогнозирования и простые модели, и другие распространенные методы и модели наукастинга, такие как MIDAS и Bridge-уравнения. Особенно подчеркивается важность включения переменных по США, таких как занятость и бизнес-опросы, в связи с их оперативной публикацией и значительным влиянием экономики США на Канаду. В целом статья подтверждает эффективность DFM в условиях разнородных и несбалансированных данных, что делает ее особенно полезной для центральных банков и аналитиков, занимающихся наукастингом и краткосрочным прогнозированием.

Несмотря на то что факторные модели получили широкое распространение в задачах наукастинга благодаря способности обрабатывать большие массивы разнородных данных, они обладают рядом ограничений. Во-первых, такие модели требуют сложной процедуры спецификации и оценки, в том числе выбор числа факторов и лагов, что может приводить к переобучению и высокой вычислительной нагрузке (Doz et al., 2011; BańBura et al., 2010). Во-вторых, интерпретация латентных факторов затруднена, что снижает прозрачность модели для прикладного использования в центральных банках (Giannone et al., 2008). Кроме того, несмотря на способность обрабатывать несбалансированные панели, эффективность DFM может значительно снижаться при большом количестве пропусков или высокой гетерогенности данных (Mariano and Murasawa, 2003). Кроме того, модели слабо адаптируются к структурным сдвигам в экономике и требуют регулярной перекалибровки. Наконец, они чувствительны к ошибкам в предварительной обработке данных и подвержены влиянию публикационных лагов и последующих уточнений данных (Angelini et al., 2011). Все это указывает на необходимость осторожного

использования факторных моделей и подкрепления их альтернативными подходами в рамках системного наукастинга.

Еще один класс моделей, получивший популярность в последние два десятилетия и позволяющий эффективно комбинировать данные разной частоты одновременно, — это MIDAS. Идея использовать их впервые предложена в работе [Ghysels et al. \(2004\)](#). В отличие от других традиционных моделей, MIDAS позволяют напрямую включать высокочастотные индикаторы (например, месячные или даже ежедневные данные) в прогнозы низкочастотных переменных, таких как квартальный ВВП. Это делает MIDAS особенно привлекательными в контексте наукастинга, где важна оперативность и гибкость в использовании поступающей информации ([Ghysels et al., 2007](#)). Важными преимуществами MIDAS являются сокращение потери информации из-за агрегации данных, возможность осуществлять корректировку прогноза в режиме реального времени, а также простота по сравнению с DFM при сохранении высокой прогностической точности. В прикладных исследованиях показано, что MIDAS-модели часто превосходят Bridge-уравнения и другие базовые модели по своим прогностическим свойствам, особенно на ранних этапах квартала ([Clements and Galvao, 2008](#); [Marcellino and Schumacher, 2010](#)). Однако эффективность MIDAS зависит от корректного выбора спецификации весовой функции и числа включенных лагов, что требует осторожного эмпирического тестирования.

К настоящему времени в литературе по наукастингу сформировался набор усовершенствованных версий эконометрических моделей, предназначенных для работы с данными различной частоты, в том числе: (1) одно- и многомерные регрессионные модели по смешанным данным с параметрическими ограничениями (MIDAS) и без таковых (Unrestricted MIDAS — U-MIDAS); (2) модели векторной авторегрессии по данным смешанной частоты (Mixed-Frequency VAR — MF-VAR), включая их байесовские модификации (MF-BVAR); (3) DFM, адаптированные к работе с разной частотой наблюдений; (4) модели векторной авторегрессии с марковскими переключениями состояний (Markov Switching MF-VAR), применяемые к данным смешанной частоты. Эти и другие подходы систематизированы и обобщены в работах [Forni и Marcellino \(2013\)](#), а также в более поздних региональных и страновых обзорах [Linzenich and Meunier \(2024\)](#), [Жемкова \(2021\)](#), [Малюгина \(2024\)](#), отражающих актуальные прикладные практики.

В последние годы методы и алгоритмы машинного обучения все активнее используются в задачах наукастинга макроэкономических показателей, демонстрируя конкурентоспособные результаты по сравнению с традиционными эконометрическими моделями. Их ключевое преимущество заключается в способности выявлять нелинейные зависимости, эффективно обрабатывать большие и высокочастотные массивы данных, адаптироваться к структурным изменениям в экономике. Работы [Apaydin et al. \(2019\)](#) и [Coulombe et al. \(2020\)](#) показывают, что алгоритмы случайных лесов, градиентного бустинга и нейронные сети обеспечивают высокую точность краткосрочного прогнозирования ВВП и инфляции в условиях нестабильной экономической среды. [Medeiros et al. \(2021\)](#) демонстрируют, что регуляризованные методы, такие как LASSO и Elastic Net, особенно эффективны в условиях высокой размерности и коррелированности макроиндикаторов.

[Zhang et al. \(2023\)](#) сравнили эффективность различных алгоритмов машинного обучения, факторных моделей и моделей MIDAS для наукастинга темпов роста реального ВВП Китая, используя 89 макроэкономических переменных с 1995 по 2020 год. Авторы установили, что некоторые методы машинного обучения превосходят эталонную DFM.

Наибольшую эффективность продемонстрировала Ridge-регрессия, которая превзошла все остальные модели не только в отношении ошибки прогноза, но и в эффективном распознавании влияния глобального финансового кризиса и шоков COVID-19.

В исследовании [Sharma and Kathuria \(2024\)](#) отмечается, что основное преимущество традиционных эконометрических моделей для наукастинга заключается в их способности справляться с такими проблемами данных, как высокая размерность, задержка и несинхронность публикаций, частые пересмотры. Вместе с тем авторы утверждают, что традиционные модели могут оказаться непригодными при использовании больших данных, в которых наблюдается экспоненциальный рост гиперпараметров. В таком случае модели машинного обучения являются дополнительным инструментом, поскольку они позволяют работать с большими данными и проводить более качественную кросс-валидацию для эффективного выбора гиперпараметров.

[Kapetanios and Papailias \(2018\)](#) рассматривают потенциал использования больших данных в задачах макроэкономического наукастинга и проводят эмпирическое сравнение различных методов обработки высокоразмерных массивов данных. Авторы показывают, что подходы, основанные на регуляризации и машинном обучении (в частности, LASSO и Ridge-регрессия), могут существенно улучшать точность прогнозов по сравнению с традиционными моделями, особенно при наличии большого числа предикторов. Основное внимание в работе уделяется вопросам отбора переменных, устойчивости моделей к шуму и переобучению и преимуществам комбинирования сигналов из различных источников данных. Кроме того, [Kapetanios and Papailias \(2018\)](#) отмечают эффективность использования комбинированных прогнозов. Статья подчеркивает важность адаптации существующих макроэконометрических инструментов к новым реалиям, связанным с использованием больших данных.

В то же время, как подчеркивают [Rossi and Sekhposyan \(2019\)](#), важным направлением остается интеграция машинного обучения с традиционными макроэкономическими моделями — для повышения интерпретируемости результатов и соблюдения теоретических ограничений. Таким образом, машинное обучение становится неотъемлемой частью современного арсенала наукастинга, особенно в условиях цифровизации и роста доступности больших данных.

Современная макроэкономическая среда, характеризующаяся высокой волатильностью и быстрой сменой деловых циклов, требует от центральных банков, других институтов, принимающих экономические решения, а также исследователей постоянного совершенствования инструментов наукастинга. Обзор существующих подходов показывает, что традиционные модели, такие как Bridge-уравнения, DFM и MIDAS, по-прежнему актуальны и широко применяются. Бурное развитие цифровых технологий и рост доступности высокочастотных и неструктурированных данных открывают новые возможности для повышения точности краткосрочных прогнозов. В этих условиях методы машинного обучения становятся важным дополнением к классическим инструментам, особенно в задачах отбора признаков, выявления нелинейностей и адаптации к структурным сдвигам. Необходимость дальнейших теоретических и прикладных исследований, направленных на интеграцию этих методов в стандартный прогнозный арсенал, становится все более очевидной. Развитие гибридных моделей и систематическая оценка их прогностической эффективности — ключевые направления для дальнейших научных и прикладных исследований в области наукастинга.

2. Обзор моделей

Приведем описание сравниваемых в настоящей работе традиционных эконометрических моделей для наукастинга, таких как Bridge-уравнение, MIDAS и факторная модель, а также методов и алгоритмов машинного обучения.

Традиционные эконометрические модели

Bridge-уравнение

В качестве Bridge-уравнения мы применяем авторегрессионную модель с распределенными лагами (ARDL-модель), включающую лаговые значения зависимой переменной, текущее и лаговые значения высокочастотной объясняющей переменной. Модель имеет следующий вид:

$$y_{t_q} = a_0 + \sum_{i=1}^p a_i L^i y_{t_q} + \sum_{j=0}^q b_j L^j x_{t_q} + \varepsilon_{t_q} \quad (1)$$

где y_{t_q} — низкочастотный временной ряд (зависимая переменная),
 x_{t_q} — высокочастотный временной ряд (объясняющая переменная),
 a_i, b_j — неизвестные параметры модели, оцениваемые по МНК.

Уравнение (1) называется Bridge-уравнением, поскольку оно устанавливает связь между высокочастотными и низкочастотными данными. Это позволяет оценить (прогнозировать) значения низкочастотной переменной на основе имеющихся на данный момент высокочастотных данных. Прогнозирование с использованием модели (1) осуществляется в два этапа. На первом этапе высокочастотный ряд прогнозируется до окончания текущего квартала, после чего полученные значения агрегируются до квартальной частоты. Для прогнозирования высокочастотного ряда применяется традиционная модель ARIMA. На втором этапе агрегированные значения высокочастотной переменной подставляются в модель (1) вместо x_{t_q} , и на основе оцененных параметров модели рассчитывается прогноз y_{t_q} . Отметим, что оптимальные порядки лагов определяются с помощью информационного критерия Акаике (AIC). В случае использования нескольких высокочастотных индикаторов итоговый прогноз y_{t_q} формируется как среднее арифметическое прогнозов, полученных по разным Bridge-уравнениям.

В рамках данного исследования подход на основе Bridge-уравнения будет служить эталоном при сравнении с альтернативными подходами для наукастинга.

Модель MIDAS (Mixed Data Sampling)

Отличительной особенностью модели MIDAS является то, как в ней манипулируют высокочастотными данными. В MIDAS для связывания высоко- и низкочастотных данных применяют полиномиальные весовые функции. Такой подход делает MIDAS инструментом прямого прогнозирования. Это ее основное отличие от Bridge-уравнений, где каждый высокочастотный показатель прогнозируется отдельно, а потом результаты применяются

для наукастинга. В MIDAS значения низкочастотной переменной (квартальный прирост реального ВВП или дефлятор ВВП) напрямую связываются с текущим и лаговыми значениями высокочастотного показателя (например, индекс промышленного производства).

Для более детального объяснения, каким образом связываются низко- и высокочастотные переменные, сделаем следующие обозначения:

$t = 1, 2, \dots, T$ — это временные такты для низкочастотного показателя (y_t^L),

$t = 1, 2, \dots, m \cdot T$ — это временные такты для высокочастотного показателя (x_t^H),

m — целое число, которое показывает, сколько раз появляется значение высокочастотной переменной за единицу времени низкочастотного показателя.

Так, например, для квартального прироста ВВП и месячного индекса промышленного производства $m = 3$.

Низкочастотную переменную (например, темп прироста ВВП) обозначим y_t^L , высокочастотную переменную (индекс промышленного производства) обозначим x_t^H . Тогда, если мы хотим построить регрессионную модель с текущим значением и первым лагом x_t^H , матрица данных в MIDAS примет следующий вид:

$$\begin{bmatrix} y_2^L \\ y_3^L \\ \vdots \\ y_n^L \end{bmatrix} = \begin{bmatrix} x_6^H & x_5^H & \dots & x_1^H \\ x_9^H & x_8^H & \dots & x_4^H \\ \dots & \dots & \dots & \dots \\ x_{3n}^H & x_{3n-1}^H & \dots & x_{3n-5}^H \end{bmatrix} \times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_5 \end{bmatrix} + \begin{bmatrix} \epsilon_2 \\ \epsilon_3 \\ \vdots \\ \epsilon_n \end{bmatrix} \quad (2)$$

Модель включает текущее значение и первый лаг высокочастотной переменной, поэтому второй квартал низкочастотной переменной y_2^L будет зависеть от трех текущих значений x_t^H , которым соответствуют $\{x_6^H, x_5^H, x_4^H\}$, и трех значений высокочастотной переменной за предыдущий квартал, которым соответствуют $\{x_3^H, x_2^H, x_1^H\}$. Аналогичным образом объясняются и все остальные строки приведенной модели.

Следующий шаг — оценивание модели MIDAS. Базовую модель MIDAS, в которой применяется один высокочастотный показатель, можно представить следующим образом:

$$y_{t+h}^L = a_h + b_h C \left(L^{\frac{1}{m}}; \theta \right) x_t^H + \epsilon_{t+h}^L, \quad (3)$$

где a_h, b_h — неизвестные параметры модели, которые надо оценить;

$C \left(L^{\frac{1}{m}}; \theta \right) = \sum_{i=0}^N c(i; \theta) L^{i/m}$ — коэффициенты полинома степени N ;

$L^{i/m} x_t = x_{t-i/m}$ — лаговый оператор.

В развернутой форме модель (3) можно записать как:

$$y_{t+h}^L = a_h + b_h \left(c(0; \theta) x_{t-0/m} + c(1; \theta) x_{t-1/m} + \dots + c(N; \theta) x_{t-N/m} \right) + \epsilon_{t+h}^L. \quad (4)$$

Как видно из модели (4), оценивание параметров модели (3) реализуется в два этапа. На первом этапе рассчитываются значения весов $c(i; \theta)$. На втором этапе, применяя рассчитанные веса лаговых значений, x_t агрегируются, в результате чего формируется взвешенная переменная. Смысл применения весовых коэффициентов состоит в том, чтобы более поздним лаговым значениям x_t^H придать меньший вес и наоборот. Весовые коэффициенты $c(i; \theta)$ рассчитываются по формуле:

$$c(i; \theta) = \frac{x_i^{\theta_1 - 1} (1 - x_i)^{\theta_2 - 1}}{\sum_{i=1}^N x_i^{\theta_1 - 1} (1 - x_i)^{\theta_2 - 1}}, \quad (5)$$

где $x_i = \frac{i}{N+1}$, а весовые коэффициенты $c(i; \theta)$ зависят от θ_1 и θ_2 .

Задача состоит в том, чтобы подобрать такие значения θ_1 и θ_2 , при которых оцененная модель (3) будет иметь минимальное значение суммы квадратов ошибок. Такая задача может быть решена на основе нелинейного метода наименьших квадратов. Другими словами, задача в том, чтобы одновременно найти такие значения $a_h, b_h, \theta_1, \theta_2$, которые минимизируют сумму квадратов ошибок в модели (3).

В работе Ghysels et al. (2004) демонстрируется, что модель MIDAS позволяет получить более эффективные оценки по сравнению с подходом, когда высокочастотный ряд агрегируется к более низкочастотному ряду, а потом делается оценка модели. Одновременно авторы отмечают, что смещение оценок в результате дискретизации объясняющей переменной почти такое же, как и в случае с обычной моделью с распределенными лагами. Причем смещение оценок убывает при увеличении частотности объясняющих переменных.

Факторная модель

В отличие от Bridge-уравнения и модели MIDAS, данный подход основан на использовании сравнительно большой базы исходных данных. Поясним наукастинг низкочастотной целевой переменной с применением факторной модели с помощью методологии, предложенной Doz et al. (2011). Это методология разработана для извлечения небольшого количества динамических главных компонент (факторов) из крупного массива данных, что позволяет эффективно сжимать информацию и учитывать скрытые зависимости.

Методология Doz et al. (2011) предусматривает два этапа. На первом этапе рассчитываются статические главные компоненты на основе сбалансированных данных. Эти компоненты затем используются для определения начальных условий фильтра Калмана. На втором этапе, используя полученные начальные условия, применяются фильтр Калмана и техника обратного сглаживания для расчета динамических главных компонент.

Модель имеет следующий вид:

$$x_{t_m} = \Lambda f_{t_m} + \xi_{t_m}, \quad \xi_{t_m} \sim N(0, \Sigma_\xi), \quad (6)$$

$$f_{t_m} = \sum_{i=1}^p A_i f_{t_m-i} + B \eta_{t_m}, \quad \eta_{t_m} \sim N(0, I_q). \quad (7)$$

Уравнение (6) связывает N месячных показателей с $r \times 1$ главными компонентами через матрицу факторных нагрузок Λ .

Уравнение (7) описывает правило, по которому генерируется ненаблюдаемая компонента f_{t_m} , где A — матрица коэффициентов, оцененная на основе статических главных компонент с использованием МНК.

Прогноз зависимой переменной на текущий квартал осуществляется на основе Bridge-уравнения по формуле:

$$\hat{y}_{t_q} = \alpha + \beta \hat{f}_{t_q,r} \quad (8)$$

где \hat{f}_{t_q} — агрегированные до квартальной частоты значения \hat{f}_{t_m} .

Методы и алгоритмы машинного обучения

Общая черта подходов к машинному обучению заключается в том, что алгоритмы реализуются с минимальным вмешательством человека. Один из вариантов таксономии, используемой для классификации алгоритмов машинного обучения, представлен в рабочем документе (Цукарев, Погосян, Лемба, 2024).

В настоящем исследовании коротко приведены те методы и алгоритмы машинного обучения, которые являются достаточно популярными для наукастинга и которые можно использовать для выявления сравнительных преимуществ.

Ridge-регрессия (гребневая регрессия)

Алгоритм гребневой регрессии схож с МНК. Он разработан для преодоления нестабильности оценок МНК² путем пенализации суммы квадратов параметров на основе регуляризации³ $L_2 \left(\sum_{j=1}^p \beta_j^2 \right)$. В частности, коэффициенты модели оцениваются по следующей оптимизационной формуле (Hoerl and Kennard, 1970):

$$\hat{\beta} = \operatorname{argmin}_{\beta} \left[\sum_{i=1}^n \underbrace{\left(y_i - \sum_{j=1}^p x_{ij} \beta_j \right)^2}_{RSS} + \underbrace{\lambda \sum_{j=1}^p \beta_j^2}_{\text{Ridge-пенализация}} \right], \quad (9)$$

где n — количество наблюдений,
 p — количество объясняющих переменных,
 $\lambda \geq 0$ — параметр, который контролирует величину сжатия⁴.

² При высокой корреляции между независимыми переменными (мультиколлинеарности).

³ 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.

⁴ Чем больше значение λ , тем больше величина сжатия.

Перепишем формулу (9) в матричном виде:

$$RSS(\lambda) = (Y - X\beta)^T(Y - X\beta) + \lambda\beta^T\beta. \quad (10)$$

Взяв производные по β , получим оценки неизвестных параметров:

$$\hat{\beta} = (X^T X + \lambda I)^{-1} X^T Y, \quad (11)$$

где I — единичная матрица.

Согласно формуле (11), в отличие от МНК, к диагонали матрицы $X^T X$ добавляется положительная константа λ . Если $\lambda = 0$, то оценки Ridge-регрессии будут в точности равны результатам МНК, а если $\lambda \rightarrow \infty$, то коэффициенты модели стремятся к нулю. Таким образом, на практике параметр λ изменяется между 0 и ∞ .

Параметры Ridge-регрессии можно оценить с помощью указанной выше аналитической формулы. Однако для ускорения вычислений можно использовать и алгоритм градиентного спуска (*gradient descent*). 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):

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left[\underbrace{\sum_{i=1}^n \left(y_i - \sum_{j=1}^p x_{ij} \beta_j \right)^2}_{RSS} + \underbrace{\lambda \sum_{j=1}^p |\beta_j|}_{LASSO - \text{пенализация}} \right]. \quad (12)$$

Ограничение $\sum_{j=1}^p |\beta_j|$ делает решение нелинейным, поэтому применяют алгоритм градиентного спуска.

Elastic Net (Эластичная сеть)

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

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left[\sum_{i=1}^n \left(y_i - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p \left[\underbrace{(1-a)\beta_j^2}_{Ridge} + a \underbrace{|\beta_j|}_{LASSO} \right] \right], \quad (13)$$

где a — 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:

$$RSS_m = \sum_{n \in N_m} (y_n - \bar{y}_m)^2, \quad (14)$$

где y_n — наблюдения в узле N_m .

Функцией потерь для всего дерева решений является 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:

$$RSS_T = \sum_m \sum_{n \in N_m} I_m RSS_m. \quad (15)$$

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:

$$\Delta RSS_T = RSS_m - (RSS_{C_m^L} + RSS_{C_m^R}). \quad (16)$$

Как видим, ΔRSS_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-повторений⁵, максимальная и минимальная глубина дерева решений⁶ (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 (Цукарев, Погосян, Лемба, 2024).

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 лишь некоторые из случайно отобранных переменных (Breiman, 2001; Цукарев, Погосян, Лемба, 2024).

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 $n = 1, 2, \dots, N$:

$$L_n^t = \frac{|y_n - \hat{y}_n|}{D^t}, \quad (17)$$

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

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

$$\bar{L}^t = \sum_{n=1}^N L_n^t w_n^t. \quad (18)$$

⁵ Как правило, около 100 бутстрэп-повторений достаточно для получения стабильных прогнозов дерева решений.

⁶ Максимальную глубину задают таким образом, чтобы не произошло переобучения дерева решений, а минимальный размер обычно равен 2.

Если $\bar{L}_n^t \geq 0,5$, то итерационный процесс завершается, и общее количество бутстрэп-повторений $T = 100$ заменяется на $t - 1$, где $t = 1, 2, \dots, T$. В обратном случае итерационный процесс продолжается.

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

6. Обновляем весовые коэффициенты модели по 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 $\ln\left(\frac{1}{\beta^t}\right)$.

Support Vector Machine (метод опорных векторов)

Метод опорных векторов — это алгоритм машинного обучения, изначально разработанный для задач классификации, но эффективно применяемый и в регрессии. Базовая методология SVM представлена в работе [Cortes and Vapnik \(1995\)](#). Основная идея данного метода заключается в том, чтобы найти такую гиперплоскость, которая максимально разделяет наблюдения разных классов (в случае классификации) или аппроксимирует зависимость между переменными с минимальной ошибкой (в случае регрессии).

Ключевая особенность метода — использование ядерных функций (*kernel functions*), позволяющих эффективно моделировать нелинейные зависимости путем преобразования исходных данных в пространство более высокой размерности, где линейное разделение становится возможным. При этом модель сохраняет хорошую обобщающую способность за счет структурного контроля над ошибками (*regularization*) и выбора допустимой зоны отклонений (ϵ -insensitive zone for SVR).

В контексте макроэкономического наукастинга SVM применяется для прогнозирования ключевых показателей, таких как ВВП, инфляция или индекс деловой активности, особенно когда присутствуют нелинейные зависимости или сложные взаимодействия между предикторами ([Huang et al., 2005](#); [Kim, 2003](#); [Tay and Cao, 2001](#)). Метод демонстрирует устойчивость к переобучению, особенно на небольших выборках, что делает его полезным в условиях ограниченного количества квартальных наблюдений.

SVM предполагает минимизацию $J(\beta) = \frac{1}{2} \beta' \beta$ при ограничении $\forall n : |y_n - (x_n' \beta + b)| \leq \epsilon$.

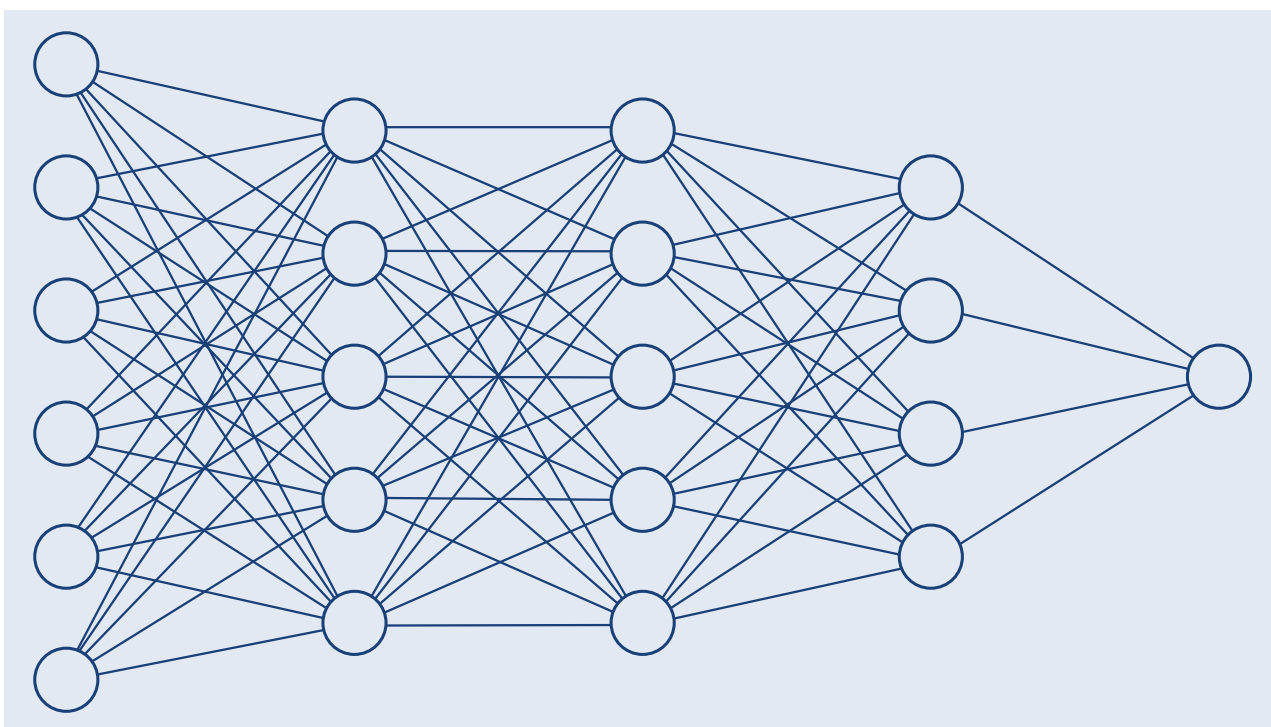
Оптимальные значения β , минимизирующие $J(\beta)$, находят с помощью алгоритма градиентного спуска. Причем те ошибки, которые находятся на расстоянии ϵ от значения y , игнорируются и приравниваются к нулю. Вышесказанное можно описать следующим образом:

$$L_\epsilon = \begin{cases} 0, & \text{если } |y - f(x)| \leq \epsilon \\ |y - f(x)| - \epsilon, & \text{в остальных случаях.} \end{cases} \quad (19)$$

Multilayer Perceptron (многослойный перцептрон)

Многослойный перцептрон (многослойная нейронная сеть) является обобщением однослойной нейронной сети. Каждый слой состоит из нейронов, причем количество нейронов первого слоя соответствует количеству входных переменных $x = (x_1, x_2, \dots, x_n)$. Последующие слои называются скрытыми (*hidden layers*). Отдельные нейроны каждого слоя представляют собой линейную комбинацию нейронов предыдущего слоя. Например, на [Рисунке 1](#) изображена схема нейронной сети с тремя скрытыми слоями.

Рисунок 1. Архитектура нейронной сети



Входной слой (см. [Рисунок 1](#)) содержит шесть нейронов, так как количество входных переменных равно 6. Далее следуют три скрытых слоя, количество нейронов у первых двух слоев равно 5, а у третьего скрытого слоя — 4. Выходной слой состоит из одного нейрона.

Алгоритм взаимодействия между нейронами работает следующим образом:

1. В первом слое для каждого нейрона в зависимости от количества предикторов случайным образом генерируются весовые коэффициенты.
2. Применяя исходную матрицу весов, рассчитываются выходные переменные, на основе которых оптимизируются весовые коэффициенты.
3. Оптимизированные матрицы весовых коэффициентов передаются в следующий слой, для которых они служат исходными весами.
4. Оптимизация весовых коэффициентов на каждом слое осуществляется на основе алгоритма обратного распространения ошибки (*backpropagation*).

В случае [Рисунка 1](#), после третьего слоя, где количество нейронов равно четырем, алгоритм завершается и делается прогноз. Подробно этот алгоритм рассматривается в [Chong and Zak \(2013\)](#), [Цукарев, Погосян, Лемба \(2024\)](#)⁷.

Recurrent Neural Network (рекуррентная нейронная сеть)

Рекуррентная нейронная сеть — это класс нейронных сетей для обработки последовательностей данных, таких как временные ряды. В отличие от обычных нейросетей, RNN обладают внутренней памятью, что позволяет им учитывать зависимость текущего значения от предыдущих состояний. Ключевая идея RNN заключается в циклической передаче информации: на каждом шаге модель получает входное значение и одновременно учитывает скрытое состояние, накопленное из предыдущих шагов. Это делает RNN особенно подходящими для задач прогнозирования в экономике, где наблюдается временная зависимость между показателями (например, между текущим и прошлыми значениями ВВП, инфляции, производственной активности и т. д.).

Однако классические RNN сталкиваются с проблемами затухающего или взрывающегося градиента, из-за чего плохо справляются с длинными последовательностями. Для решения проблемы разработаны улучшенные архитектуры, такие как LSTM и GRU, которые позволяют моделировать более долгосрочные зависимости и широко применяются в экономических задачах наукастинга и прогноза временных рядов (например, в работах [Bandara et al., 2020](#); [Fischer and Krauss, 2018](#)).

В целом алгоритм RNN можно описать следующими шагами⁸:

1. Начало каждой последовательности инициализируется с помощью нулей.
2. Из всей входной последовательности извлекается фрагмент последовательности данных и целевых значений.
3. Для текущей последовательности применяется метод прямого распространения и рассчитываются выходные вероятности.
4. На основе метода обратного распространения обновляются весовые коэффициенты, с помощью которых рассчитываются градиенты параметров.
5. Рассчитываются и обновляются функция потерь посредством обновленных весов.
6. Обновленные весовые коэффициенты предыдущего скрытого состояния применяются для следующей последовательности данных.

⁷ Для читателя также может быть полезным репозиторий с обучающим пособием, в котором подробно демонстрируется информация о многослойном перцептрон для задачи регрессии: https://github.com/AndreasLeitherer/Tutorial_multilayer_perceptron/blob/main/nn_regression.ipynb.

⁸ Подробное описание алгоритма RNN и соответствующих функций, а также коды для Python можно найти по ссылке: <https://www.kaggle.com/code/fareselmenshawii/rnn-from-scratch>.

3. Данные и программное обеспечение

В рамках настоящего исследования были использованы временные ряды макроэкономических показателей по Республике Армения и Республике Беларусь за период с 2002 по 2024 год. Основными целевыми переменными выступили темпы прироста реального ВВП и дефлятор ВВП, имеющие квартальную частотность. Все остальные показатели применялись с месячной частотностью. Полный перечень анализируемых индикаторов приведен в [Приложениях 1 и 2](#).

Для Армении было отобрано 22 высокочастотных макроэкономических показателя, для Беларуси — 20. В выборку вошли индикаторы, характеризующие различные секторы экономики. Реальный сектор представлен показателями объемов выпуска и отраслевыми ценами производителей. Внешнеэкономический сектор включает данные по экспорту и импорту товаров и услуг, валовым международным резервам и номинальному обменному курсу. Финансовый сектор охватывает ключевые денежно-кредитные и финансовые индикаторы.

Следует отметить, что количество объясняющих показателей могло быть больше, однако нет какой-либо доказательной основы того, что при увеличении количества переменных прогноз улучшится ([Barhoumi et al., 2010](#); [Alvarez and Perez-Quiros 2016](#)).

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 процентных ставок,
3. приведение временных рядов к стационарному виду посредством взятия первых разностей.

Эконометрическое моделирование (за исключением Bridge-уравнения) выполнялось в среде MATLAB⁹. Интересное практическое решение было также предложено [Linzenich and Meunier \(2024\)](#), которые подготовили программный пакет для наукастинга при помощи Bridge-уравнения, MIDAS и DFM¹⁰.

Алгоритмы машинного обучения были реализованы на Python с использованием библиотек, включая scikit-learn для классических методов, Keras и TensorFlow — для задач глубокого обучения. Дополнительно авторами данного исследования разработаны собственные вспомогательные коды для выполнения расчетов и визуализации данных.

⁹ MATLAB-коды для MIDAS можно найти по ссылке: <https://www.mathworks.com/matlabcentral/fileexchange/45150-midas-matlab-toolbox>. MATLAB-коды для DFM располагаются здесь: <https://github.com/FRBNY-TimeSeriesAnalysis/Nowcasting>.

¹⁰ Пакет можно найти по ссылке: https://github.com/baptiste-meunier/Nowcasting_toolbox.

4. Результаты экспериментов

Основная задача экспериментов в рамках данного исследования заключалась в оценке темпов прироста реального ВВП и дефлятора ВВП посредством актуальных месячных макроэкономических показателей.

Стоит отметить, что ключевая методологическая сложность при проведении наукастинга в режиме реального времени связана с отсутствием опубликованных высокочастотных данных за отдельные месяцы текущего квартала. Для решения проблемы предлагается использовать подход, описанный в рабочем документе (Цукарев, Погосян, Лемба, 2024), где в качестве базового метода применяется модель ARIMA.

Для сравнения прогностических свойств методов наукастинга, рассмотренных в третьем разделе, был осуществлен регрессионный эксперимент. Его схема включала следующие этапы:

1. В результате приведения имеющихся высокочастотных рядов к квартальной размерности сформирована выборка данных за период с 2002Q2 по 2024Q4 (91 наблюдение для каждой переменной).
2. Первоначальное деление всей выборки на две части — обучающую (первые 70% наблюдений, с 2002Q2 по 2018Q2, 65 наблюдений) и тестируемую (оставшиеся 30% наблюдений, с 2018Q3 по 2024Q4, 26 наблюдений).
3. Оценивание модели на основе фактических данных обучающей выборки (2002Q2 — 2018Q2).
4. Осуществление наукастинга показателя на один квартал вперед (первый прогнозируемый период — 2018Q3) с использованием квартальных значений высокочастотных рядов.
5. Увеличение обучающей выборки на один квартал (с 2002Q2 по 2018Q3) и переоценка параметров модели.
6. Выполнение наукастинга следующего квартала посредством переоцененной модели (2018Q4).
7. Повторение шагов 5–6 до тех пор, пока обучающая выборка не охватила период до 2024Q3 включительно.
8. Сравнение полученных точечных наукастов с фактическими значениями и расчет двух метрик точности прогноза, которые являются функциями потерь (*loss functions*):

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

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

где y_t и \hat{y}_t — фактическое и прогнозируемое значение показателя,
 n — общее количество оценок (в данном случае $n = 26$).

Результаты расчетов RMSE и MAD для темпов прироста реального ВВП и дефлятора ВВП Армении и Беларуси представлены в [Таблицах 1–4](#).

Таблица 1. Результаты оценки точности наукастинга темпов прироста реального ВВП Армении (для периода с 2018Q3 по 2024Q4)

| Модель | RMSE | RMSE (по отнош. к Bridge) | DM-stat. | MAD | MAD (по отнош. к Bridge) | DM-stat. |
|------------------|------|---------------------------------|----------|------|--------------------------------|----------|
| Bridge-уравнение | 3,56 | – | – | 2,20 | – | – |
| MIDAS | 3,38 | 0,95 | 0,42 | 2,15 | 0,98 | 0,07 |
| DFM | 3,63 | 1,02 | 0,15 | 2,42 | 1,10 | 1,01 |
| Ridge | 3,44 | 0,96 | 0,37 | 2,28 | 1,03 | 0,53 |
| LASSO | 3,38 | 0,95 | 0,43 | 2,35 | 1,07 | 0,74 |
| Elastic Net | 3,46 | 0,97 | 0,26 | 2,42 | 1,10 | 1,03 |
| Boosting | 3,52 | 0,99 | 0,53 | 2,13 | 0,97 | 0,72 |
| Bagging | 3,59 | 0,99 | 0,08 | 2,15 | 0,98 | 1,48 |
| Random Forest | 3,41 | 0,96 | 0,08 | 2,12 | 0,96 | 1,48 |
| SVM | 3,45 | 0,97 | 0,22 | 2,53 | 1,15 | 1,11 |
| MLP | 3,45 | 0,97 | 0,28 | 2,57 | 1,17 | 1,25 |
| RNN | 3,42 | 0,96 | 0,61 | 2,16 | 0,98 | 0,07 |

Таблица 2. Результаты оценки точности наукастинга дефлятора ВВП Армении (с 2018Q3 по 2024Q4)

| Модель | RMSE | RMSE (по отнош. к Bridge) | DM-stat. | MAD | MAD (по отнош. к Bridge) | DM-stat. |
|------------------|------|---------------------------------|----------|------|--------------------------------|----------|
| Bridge-уравнение | 2,95 | – | – | 2,10 | – | – |
| MIDAS | 2,29 | 0,78 | 1,98** | 1,52 | 0,72 | 2,79*** |
| DFM | 2,65 | 0,90 | 0,81 | 1,60 | 0,76 | 2,24** |
| Ridge | 2,34 | 0,79 | 2,30** | 1,43 | 0,68 | 4,06*** |
| LASSO | 2,22 | 0,75 | 2,06** | 1,51 | 0,72 | 2,79*** |
| Elastic Net | 2,34 | 0,79 | 2,01** | 1,54 | 0,73 | 2,62*** |
| Boosting | 2,41 | 0,81 | 1,90* | 1,52 | 0,72 | 2,74*** |
| Bagging | 2,36 | 0,80 | 1,90* | 1,56 | 0,74 | 2,74*** |
| Random Forest | 2,33 | 0,79 | 1,90* | 1,52 | 0,72 | 2,74*** |
| SVM | 2,51 | 0,85 | 1,09 | 1,65 | 0,78 | 1,84* |
| MLP | 2,58 | 0,87 | 0,88 | 1,74 | 0,83 | 1,22 |
| RNN | 2,17 | 0,73 | 2,67** | 1,47 | 0,70 | 3,87*** |

*, **, *** — существенна, соответственно, на 10%, 5% и 1%-ном уровне значимости.

Таблица 3. Результаты оценки точности наукастинга темпов прироста реального ВВП Беларуси (с 2018Q3 по 2024Q4)

| Модель | RMSE | RMSE (по отнош. к Bridge) | DM-stat. | MAD | MAD (по отнош. к Bridge) | DM-stat. |
|---------------|------|---------------------------------|----------|------|--------------------------------|----------|
| Bridge | 2,34 | – | – | 2,02 | – | – |
| MIDAS | 2,12 | 0,91 | 0,81 | 1,74 | 0,86 | 0,71 |
| Factor model | 2,05 | 0,88 | 1,03 | 1,68 | 0,83 | 0,92 |
| Ridge | 1,97 | 0,84 | 1,21 | 1,54 | 0,76 | 1,46 |
| LASSO | 1,93 | 0,82 | 1,33 | 1,51 | 0,75 | 1,56 |
| Elastic Net | 1,94 | 0,83 | 1,32 | 1,52 | 0,75 | 1,53 |
| Boosting | 1,65 | 0,71 | 1,68* | 1,48 | 0,77 | 1,69* |
| Bagging | 1,75 | 0,75 | 1,55 | 1,54 | 0,76 | 1,64 |
| Random Forest | 1,66 | 0,71 | 1,66 | 1,47 | 0,73 | 1,71* |
| SVM | 1,44 | 0,62 | 1,96** | 1,43 | 0,71 | 1,98** |
| MLP | 2,09 | 0,89 | 1,04 | 1,69 | 0,83 | 0,87 |
| RNN | 2,05 | 0,87 | 0,99 | 1,65 | 0,82 | 0,90 |

*, **, *** — существенна, соответственно, на 10%, 5% и 1%-ном уровне значимости.

Таблица 4. Результаты оценки точности наукастинга дефлятора ВВП Беларуси (с 2018Q3 по 2024Q4)

| Модель | RMSE | RMSE (по отнош. к Bridge) | DM-stat. | MAD | MAD (по отнош. к Bridge) | DM-stat. |
|---------------|------|---------------------------------|----------|------|--------------------------------|----------|
| Bridge | 2,14 | – | – | 1,95 | – | – |
| MIDAS | 2,10 | 0,98 | 0,25 | 1,89 | 0,97 | 0,28 |
| Factor model | 2,06 | 0,96 | 0,20 | 1,86 | 0,95 | 0,15 |
| Ridge | 1,64 | 0,77 | 1,88* | 1,42 | 0,73 | 1,76* |
| LASSO | 1,54 | 0,72 | 1,95* | 1,36 | 0,70 | 1,98** |
| Elastic Net | 1,57 | 0,73 | 1,95** | 1,37 | 0,70 | 2,05** |
| Boosting | 1,73 | 0,81 | 1,77* | 1,46 | 0,75 | 1,71* |
| Bagging | 1,70 | 0,79 | 1,80* | 1,53 | 0,79 | 1,65* |
| Random Forest | 1,65 | 0,77 | 1,80* | 1,50 | 0,77 | 1,72* |
| SVM | 1,78 | 0,83 | 1,59 | 1,54 | 0,79 | 1,71* |
| MLP | 1,80 | 0,84 | 1,56 | 1,58 | 0,81 | 1,60 |
| RNN | 1,43 | 0,67 | 2,44** | 1,40 | 0,72 | 2,14** |

*, **, *** — существенна, соответственно, на 10%, 5% и 1%-ном уровне значимости.

Согласно данным [Таблиц 1 и 3](#), содержащим результаты наукастинга темпа прироста реального ВВП, наименьшая ошибка прогнозирования по метрике RMSE для Армении достигается при использовании LASSO-регрессии. При этом полученное значение статистически не отличается от результатов традиционных методов наукастинга. В случае

Беларуси минимальная ошибка прогноза наблюдается при применении алгоритма SVM, причем его преимущество над Bridge-регрессией является статистически значимым.

Результаты наукастинга дефлятора ВВП, представленные в [Таблицах 2 и 4](#), демонстрируют, что для обеих стран наименьшая ошибка прогнозирования достигается при использовании алгоритма RNN. Это свидетельствует о высокой эффективности данного метода при моделировании динамики ценовых показателей.

Общий вывод: алгоритмы машинного обучения значимо минимизируют ошибку наукастинга, что говорит об их сравнительной точности и целесообразности для применения при наукастинге.

Далее на [Рисунках 2–5](#) представлена динамика фактических и оцененных значений реального прироста ВВП и дефлятора ВВП для Армении и Беларуси. Полученные результаты свидетельствуют о том, что методы и алгоритмы машинного обучения более точно улавливают вектор изменения рассматриваемых показателей по сравнению с традиционными моделями наукастинга.

Следующий вопрос, требующий ответа: насколько существенны различия между результатами наукастинга, полученными с помощью алгоритмов машинного обучения, и прогнозами, построенными на основе эконометрических моделей? В качестве базовой модели для сравнения выбрано Bridge-уравнение — традиционный эконометрический метод. Его использование обусловлено простотой и сопоставимостью с подходом к наукастингу в машинном обучении.

Для оценки значимости различий в точности оценок применена статистика Диболда — Мариано ([Diebold and Mariano, 1995](#)), результаты которой представлены в [Таблицах 1–4](#).

Анализ данных показывает, что для темпов прироста реального ВВП Армении (см. [Таблица 1](#)) статистически значимых различий между эконометрическими моделями и алгоритмами машинного обучения не выявлено. Однако в случае Беларуси (см. [Таблица 3](#)) методы SVM и Boosting демонстрируют существенно более высокую точность по сравнению с Bridge-уравнением. Что касается прогнозирования дефлятора ВВП, здесь алгоритмы машинного обучения в большинстве случаев превосходят традиционные подходы — это справедливо как для Армении (см. [Таблица 2](#)), так и для Беларуси (см. [Таблица 4](#)).

Рисунок 2. Наукастинг темпов прироста реального ВВП Армении

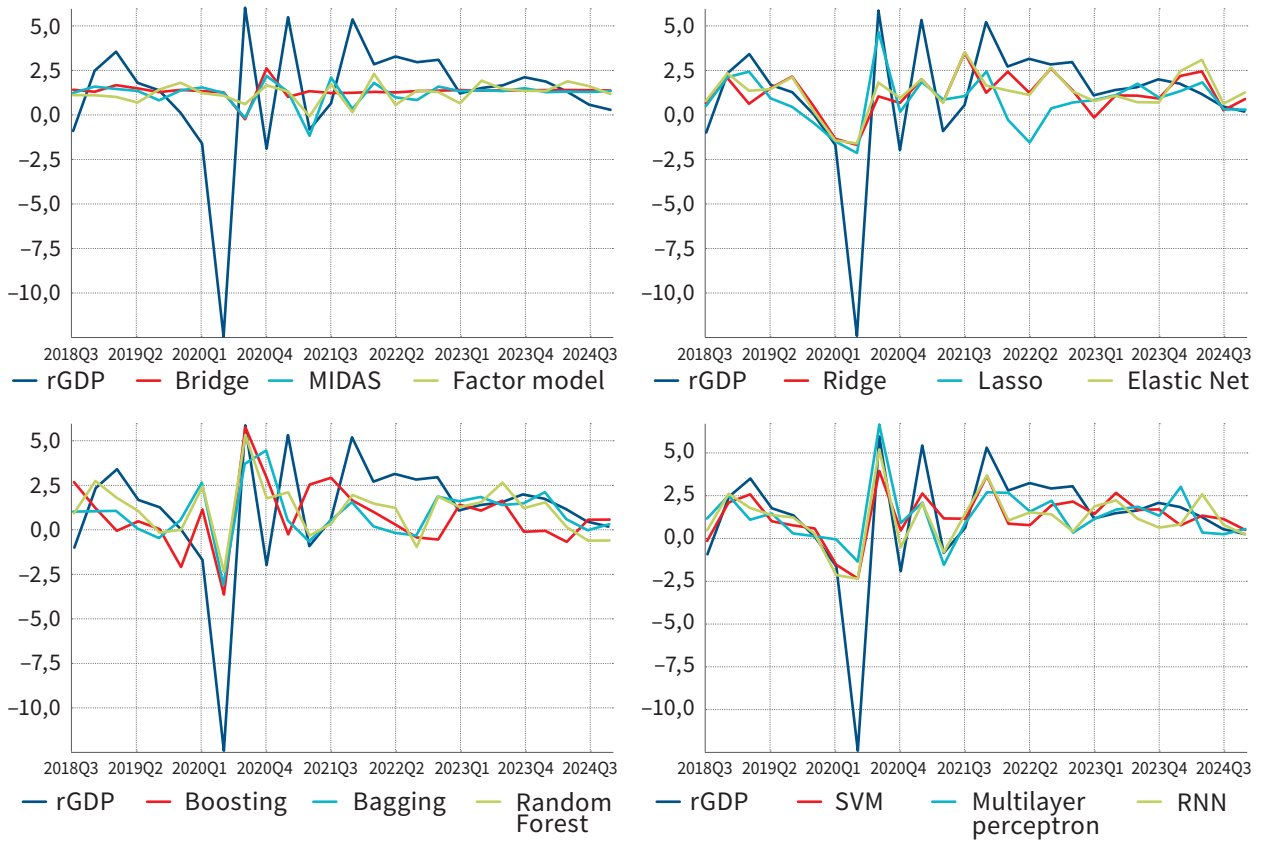


Рисунок 3. Наукастинг темпов прироста реального ВВП Беларуси

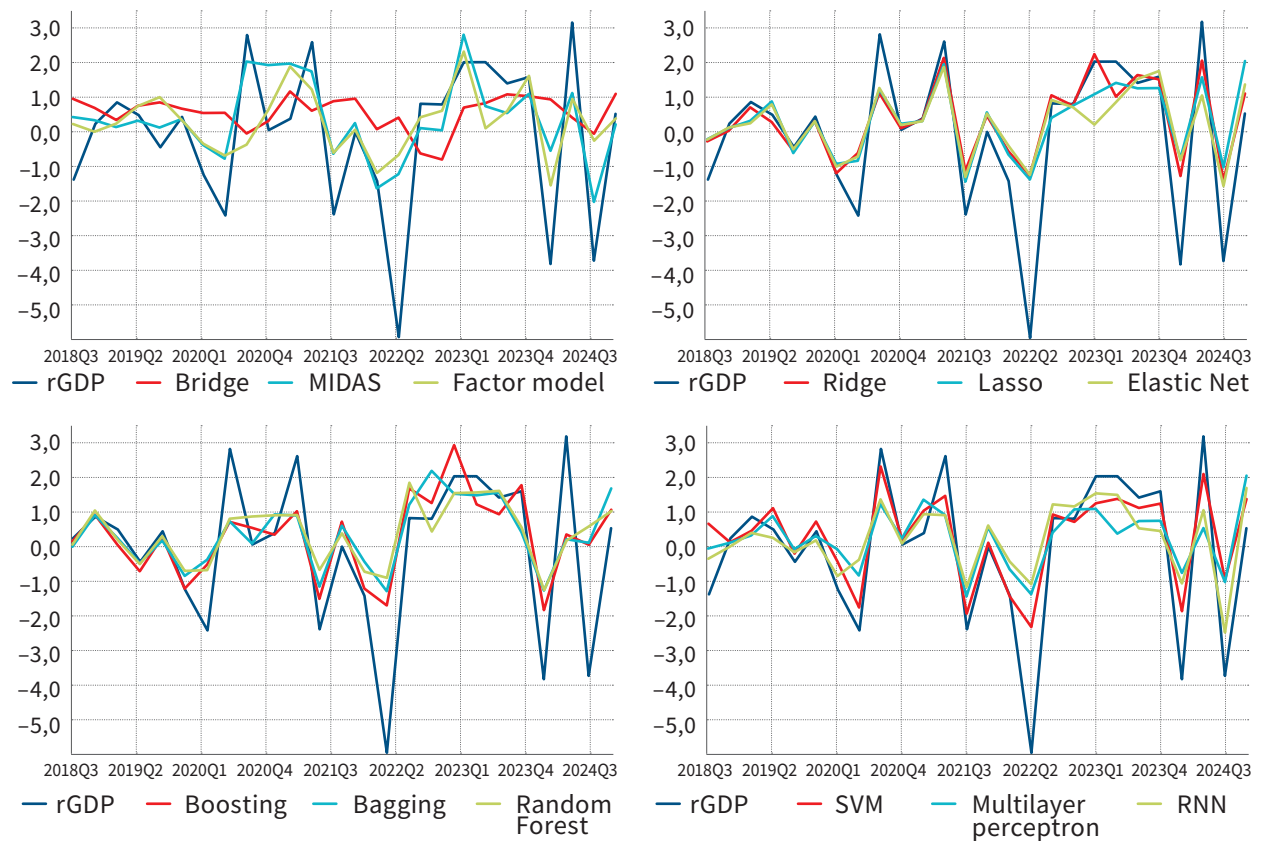


Рисунок 4. Наукастинг дефлятора ВВП Армении

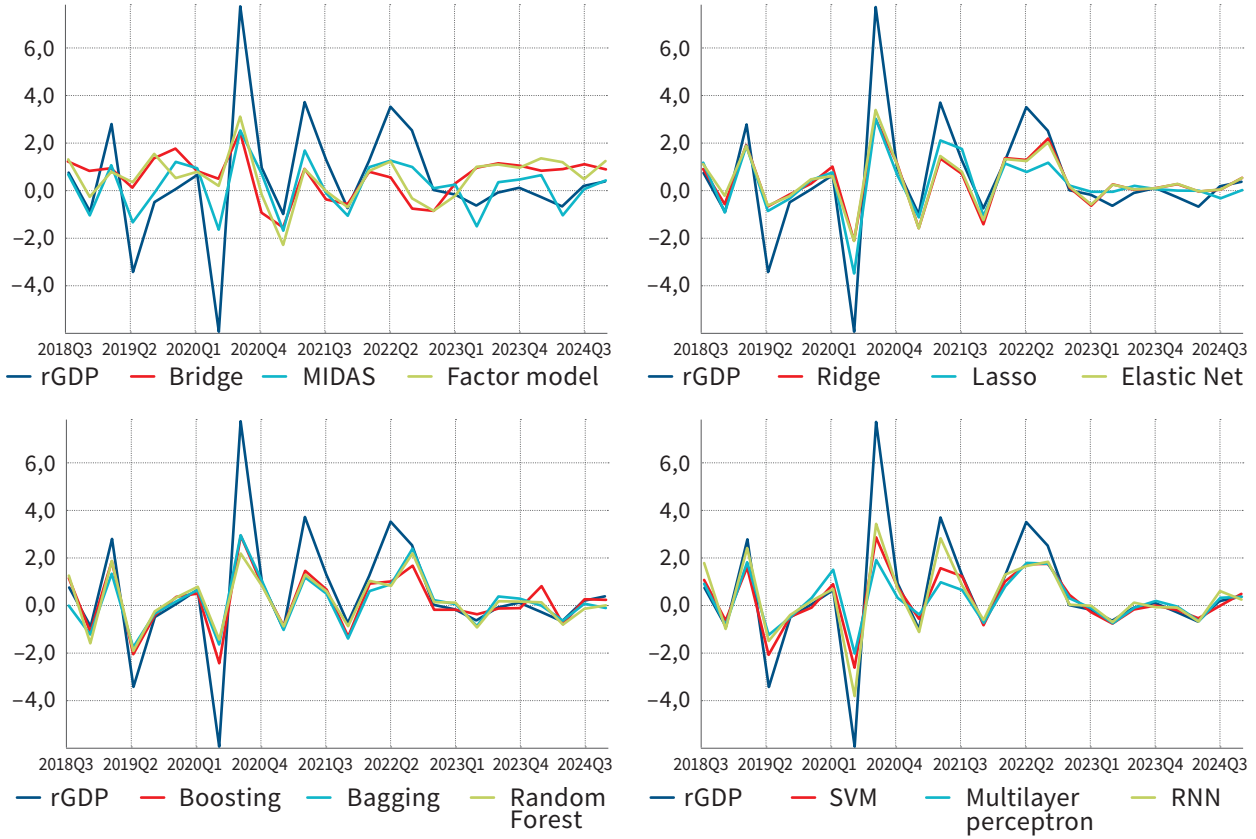
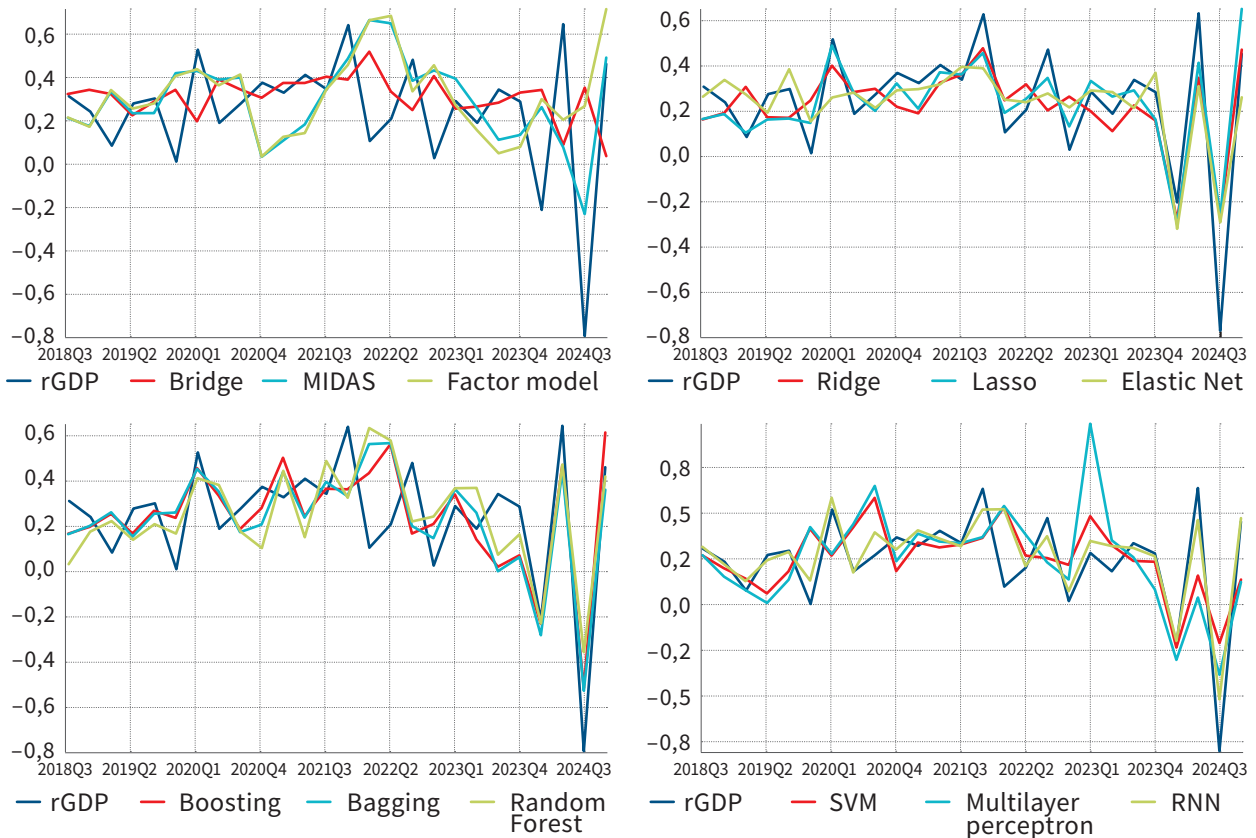


Рисунок 5. Наукастинг дефлятора ВВП Беларуси



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. Например, согласно проведенным расчетам, наилучшие результаты для темпов прироста ВВП обеспечивают LASSO и SVM, а для дефлятора ВВП — RNN. Однако в динамических условиях целесообразно регулярно переоценивать качество моделей и при необходимости переключаться между ними. Альтернативным решением может стать использование комбинированного наукастинга, позволяющего объединить преимущества разных методов.

Комбинированный наукастинг

Комбинированный наукастинг (Richardson et al., 2018) можно осуществлять путем взвешивания на основе:

1. равных весов,
2. весов, полученных на базе МНК,
3. обратных значений функции потерь,
4. обратных значений рангов функции потерь.

Согласно полученным результатам взвешивания перечисленными выше методами (Таблицы 5–8) выявлено, что все они обеспечивают более низкие значения RMSE и MAD относительно Bridge-уравнения. При этом самая высокая точность наблюдается при взвешивании с помощью МНК-весов¹¹.

Таблица 5. Комбинирование оценок темпа прироста реального ВВП Армении

| Метод взвешивания | RMSE | RMSE (по отнош. к Bridge) | MAD | MAD (по отнош. к Bridge) |
|------------------------------|------|---------------------------------|------|--------------------------------|
| Равные веса | 3,32 | 0,93 | 2,15 | 0,98 |
| МНК-веса | 2,55 | 0,72** | 1,52 | 0,69** |
| Обратные значения MSE | 3,22 | 0,91 | 2,05 | 0,93 |
| Обратные значения MSE рангов | 3,10 | 0,87 | 2,01 | 0,91 |

*, **, *** — существенна, соответственно, на 10%, 5% и 1%-ном уровне значимости.

Таблица 6. Комбинирование оценок дефлятора ВВП Армении

| Метод взвешивания | RMSE | RMSE (по отнош. к Bridge) | MAD | MAD (по отнош. к Bridge) |
|------------------------------|------|---------------------------------|------|--------------------------------|
| Равные веса | 2,11 | 0,71** | 1,50 | 0,71** |
| МНК-веса | 1,78 | 0,60*** | 1,41 | 0,67*** |
| Обратные значения MSE | 2,09 | 0,70** | 1,45 | 0,69*** |
| Обратные значения MSE рангов | 2,05 | 0,69*** | 1,47 | 0,70*** |

*, **, *** — существенна, соответственно, на 10%, 5% и 1%-ном уровне значимости.

¹¹ В данном исследовании при определении МНК-весов осуществлялось построение регрессии без константы и ограничений на параметры в соответствии с Richardson et al. (2018).

Таблица 7. Комбинирование оценок темпа прироста реального ВВП Беларуси

| Метод взвешивания | RMSE | RMSE (по отнош. к Bridge) | MAD | MAD (по отнош. к Bridge) |
|------------------------------|------|---------------------------------|------|--------------------------------|
| Равные веса | 1,75 | 0,75 [*] | 1,46 | 0,72 [*] |
| МНК-веса | 1,55 | 0,66 ^{**} | 1,33 | 0,66 ^{**} |
| Обратные значения MSE | 1,70 | 0,73 [*] | 1,45 | 0,72 |
| Обратные значения MSE рангов | 1,69 | 0,72 [*] | 1,44 | 0,71 [*] |

^{*}, ^{**}, ^{***} — существенна, соответственно, на 10%, 5% и 1%-ном уровне значимости.

Таблица 8. Комбинирование оценок дефлятора ВВП Беларуси

| Метод взвешивания | RMSE | RMSE (по отнош. к Bridge) | MAD | MAD (по отнош. к Bridge) |
|------------------------------|------|---------------------------------|------|--------------------------------|
| Равные веса | 1,35 | 0,63 ^{**} | 1,30 | 0,66 ^{**} |
| МНК-веса | 1,02 | 0,48 ^{***} | 0,99 | 0,51 ^{***} |
| Обратные значения MSE | 1,30 | 0,61 ^{**} | 1,21 | 0,62 ^{**} |
| Обратные значения MSE рангов | 1,29 | 0,60 ^{**} | 1,15 | 0,59 ^{**} |

^{*}, ^{**}, ^{***} — существенна, соответственно, на 10%, 5% и 1%-ном уровне значимости.

Основные выводы:

1. Алгоритмы машинного обучения обеспечивают более точный наукастинг по сравнению с традиционными эконометрическими моделями.
2. Согласно статистике Диболда — Мариано, оценки, полученные с помощью алгоритмов LASSO, SVM и RNN, статистически значимо отличаются от прогнозов, построенных на основе Bridge-уравнения.
3. Наилучший прогноз дает алгоритм, минимизирующий функцию потерь. Однако по мере поступления новых данных выбор модели может быть скорректирован в зависимости от динамики метрик ошибок (RMSE и MAD).
4. В качестве альтернативы можно использовать комбинированный наукастинг на основе МНК-весов. Результаты показывают, что такой подход обеспечивает статистически значимое улучшение точности по сравнению с прогнозами Bridge-уравнения.

Заключение

На текущий момент основными инструментами для наукастинга макроэкономических показателей остаются традиционные эконометрические 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.

Цель данного исследования — оценить потенциал методов машинного обучения в контексте наукастинга, а именно: 1) определить, способны ли они повысить точность прогнозирования по сравнению с традиционными подходами; 2) выяснить, могут ли эти методы не только дополнять, но и заменять классические эконометрические модели.

В рабочем документе проведено сравнительное тестирование трех традиционных эконометрических моделей (Bridge-уравнение, MIDAS и DFM) и девяти методов и алгоритмов машинного обучения (Ridge- и LASSO-регрессии, Elastic Net, Boosting, Bagging, Random Forest, SVM, MLP и RNN).

Ключевые выводы исследования:

1. Большинство рассмотренных методов и алгоритмов машинного обучения по качеству прогноза превосходят традиционные инструменты наукастинга, что подтверждается снижением значения функции потерь.
2. Согласно статистике Диболда — Мариано, наукастинг по алгоритмам LASSO, SVM и RNN статистически значимо отличается от прогноза по традиционным эконометрическим моделям.
3. Результат наукастинга существенно улучшается за счет применения комбинированного прогноза. Наибольшая точность оценок достигается при использовании комбинированного наукастинга на основе МНК-весов.

Таким образом, методы и алгоритмы машинного обучения могут рассматриваться не только как эффективное дополнение к традиционным эконометрическим подходам, но и как их альтернатива, особенно в задачах наукастинга макроэкономических показателей.

Список литературы

- Жемков, М. (2021) Краткосрочная оценка ВВП России методом комбинирования прогнозов. Серия докладов об экономических исследованиях Банка России. Доступно на: <https://www.cbr.ru/StaticHtml/File/118606/wp-feb21.pdf>.
- Малюгин, В. (2024) Краткосрочное прогнозирование и наукастинг темпов роста инфляции на основе моделей по смешанным данным. Банковский вестник, № 1/726. Доступно на: <https://www.nbrb.by/bv/pdf/articles/12119.pdf>.
- Цукарев, Т., Погосян, К., Лемба, К. (2024). Алгоритмы машинного обучения для краткосрочного прогнозирования темпов прироста реального ВВП. Рабочий документ ЕФСР №5. Доступно на: https://efsd.org/upload/iblock/d17/tbln4ev5exc6rop14b49hpwnm4l51ehc/EFSD_WP_ML_RU_2024.pdf.
- Aastveit, K. A., Gerdrup, K. R., Jore, A. S., and Thorsrud, L.A. (2011) Nowcasting GDP in real-time: A density combination approach. Norges Bank, Working Paper, No. 2011/11. Available at: <https://www.econstor.eu/bitstream/10419/209987/1/nb-wp2011-11.pdf>.
- Alvarez, R., Camacho, M. and Perez-Quiros, G. (2016). Aggregate versus disaggregate information in dynamic factor models, *International Journal of Forecasting*, vol. 32(3), pp. 680–694. Available at: <https://doi.org/10.1016/j.ijforecast.2015.10.006>.
- Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L., and Rünstler, G. (2011) Short-term forecasts of euro area GDP growth. *Econometrics Journal*, 14(1), pp. C25–C44.
- Apaydın, M., Erer, D., and Karaçor, Z. (2019) *Nowcasting GDP using machine learning algorithms: A comparative analysis*. *Journal of Business Research — Türk*, 11(4), 2587–2599.
- Baffigi, A., Golinelli, R., and Parigi, G. (2004) Bridge models to forecast the Euro area GDP, *International Journal of Forecasting*, 20(3), pp. 447–460. Available at: [https://doi.org/10.1016/S0169-2070\(03\)00067-0](https://doi.org/10.1016/S0169-2070(03)00067-0).
- BañBura, M., Giannone, D., and Reichlin, L. (2010) Nowcasting. ECB Working Paper No. 1275. Available at: <http://dx.doi.org/10.2139/ssrn.1717887>.
- Bandara, K., Bergmeir, C., and Smyl, S. (2020) Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach. *Expert Systems with Applications*, 140, 112896. DOI:10.1016/j.eswa.2019.112896.
- Barhoumi, K., Darne, O., and Ferrara, L. (2010) Are disaggregated data useful for factor analysis in forecasting French GDP? *Journal of Forecasting*, 29, pp. 132–144.
- Bencivelli, L., Marcellino, M. and Moretti, G. L., (2012) Selecting predictors by Bayesian model averaging in bridge models, Banca d'Italia WP No. 872. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2154928.
- Bernanke, B. S., Boivin, J., and Elias, P. (2005) Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. *The Quarterly journal of economics*, Vol. 120, No. 1, pp. 387–422.
- Breiman, L. (1996). Bagging predictors, *Machine Learning*, 26, pp. 123–140.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45, pp. 5–32. Available at: <https://doi.org/10.1023/A:1010933404324>.
- Breiman, L. and Ihaka, R. (1984). Nonlinear discriminant analysis via scaling and ACE. Technical Report, University of California Berkeley.
- Chernis, T., and Sekkel, R. (2017) A dynamic factor model for nowcasting Canadian GDP growth, Bank of Canada, Staff Working Paper 2017-2. DOI: <https://doi.org/10.34989/swp-2017-2>.
- Chong, E. K. P. and Zak, S. H. (2013) *An Introduction to Optimization*. 4th Edition, John Wiley & Sons, New York.
- Clements, M., and Galvao, A. (2008) Macroeconomic Forecasting with mixed frequency data: Forecasting output growth in the United States. *Journal of Business and Economic Statistics*, 26, pp. 546–554.

- Cortes, C., Vapnik, V. (1995) *Support-vector networks*. Machine Learning, 20, pp. 273–297. Available at: <https://doi.org/10.1007/BF00994018>.
- Coulombe, P. G., Leroux, M., Stevanovic, D., and Surprenant, S. (2020) *How is machine learning useful for macroeconomic forecasting?* Journal of Applied Econometrics, 37(5), pp. 847–866. Available at: <https://doi.org/10.1002/jae.2884>.
- 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), pp. 253–263.
- Diron, M. (2008): “Short-term forecasts of Euro Area real GDP growth. An assessment of real-time performance based on vintage data”, Journal of Forecasting, 27(5), pp. 371–390.
- Doz, C., Giannone, D., and Reichlin, L. (2011) A two-step estimator for large approximate dynamic factor models based on Kalman filtering. Journal of Econometrics, 164(1), pp. 188–205. Available at: <https://doi.org/10.1016/j.jeconom.2011.02.012>.
- Efron, B. and Tibshirani, R. (1993). An introduction to the Bootstrap. Chapman and Hall, London.
- Fischer, T., and Krauss, C. (2018) Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), pp. 654–669. DOI:10.1016/j.ejor.2017.11.054.
- Froni, C., and Marcellino, M. (2013) A survey of econometric methods for mixed-frequency data, Norges Bank Working Paper 2013/06. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2268912.
- 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, pp. 119–139.
- Ghysels, E., P. Santa-Clara, and Valkanov, R. (2004) The MIDAS touch: Mixed Data Sampling regression models. Mimeo, Chapel Hill, N.C. Available at: <https://escholarship.org/uc/item/9mf223rs>.
- Ghysels, E., Sinko, A., and Valkanov, R. (2007) MIDAS Regressions: Further Results and New Directions. Econometric Reviews, 26(1), pp. 53–90. Available at: <https://doi.org/10.1080/07474930600972467>
- Giannone, D., Agrippino, S.M., and Modugno, M. (2013) Nowcasting China Real GDP. Working Paper. Available at: https://cirano.qc.ca/conferences/public/pdf/realtime2013/Papers/Giannone_Miranda.pdf.
- Giannone, D., Reichlin, L., and Small, D. (2008) Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics, Vol. 55, No. 4, pp. 665–676.
- 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, pp. 55–67.
- Huang, W., Nakamori, Y., and Wang, S.Y. (2005) *Forecasting stock market movement direction with support vector machine*. Computers & Operations Research, 32(10), 2513–2522. Available at: <https://doi.org/10.1016/j.cor.2004.03.016>.
- Kapetanios, G., and Papailias, F. (2018) Big Data & Macroeconomic Nowcasting: Methodological Review, ESCoE Discussion Paper 2018-12. Available at: <https://escoe-website.s3.amazonaws.com/wp-content/uploads/2020/07/13161005/ESCoE-DP-2018-12.pdf>.
- Kim, H.Y. (2003) *Financial time series forecasting using support vector machines*. Neurocomputing, 55(1–2), pp. 307–319. Available at: [https://doi.org/10.1016/S0925-2312\(03\)00372-2](https://doi.org/10.1016/S0925-2312(03)00372-2).
- Linzenich, J., and Meunier, B. (2024) Nowcasting Made Easier: A Toolbox for Economists. ECB Working Paper No. 2024/3004, Available at: <http://dx.doi.org/10.2139/ssrn.5060436>
- Marcellino, M., and Schumacher, C. (2010) Factor-MIDAS for now- and forecasting with ragged-edge data: A model comparison for German GDP, Oxford Bulletin of Economics and Statistics, 72, pp. 518–550.

- Mariano, R. S., and Murasawa, Y. (2003) A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, 18(4), pp. 427–443.
- Matheson, T. D. (2010) An analysis of the informational content of New Zealand data releases: the importance of business opinion surveys. *Economic Modelling*, 27, pp. 304–314.
- Medeiros, M., Gabriel, F.R., Vasconcelos, G., Veiga A., and Zilberman, E. (2021) *Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods*. *Journal of Business & Economic Statistics*, Vol. 39(1), pp. 98–119. Available at: <https://doi.org/10.1080/07350015.2019.1697174>.
- 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>.
- Richardson, A., van Florenstein Mulder, T., and Vehbi, T. (2021) Nowcasting GDP using machine-learning algorithms: A real-time assessment, *International Journal of Forecasting*, Vol. 37(2), pp. 941–948. Available at: <https://doi.org/10.1016/j.ijforecast.2020.10.005>.
- Rossi, B., and Sekhposyan, T. (2019) *Alternative forecasting approaches*. In: *Handbook of Economic Forecasting*, Vol. 2B, edited by G. Elliott & A. Timmermann, Elsevier, pp. 967–1046.
- Sharma, M., Kathuria, V. (2024) Macroeconomic Nowcasting: What can Central Banks Learn from a Structured Literature Review? *Journal of Quantitative Economics*. Available at: <https://doi.org/10.1007/s40953-024-00421-x>.
- Tay, F. E. H., and Cao, L. (2001) *Application of support vector machines in financial time series forecasting*. *Omega*, 29(4), pp. 309–317. Available at: [https://doi.org/10.1016/S0305-0483\(01\)00026-3](https://doi.org/10.1016/S0305-0483(01)00026-3).
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society, Series B*, 58, pp. 267–288.
- Zhang, Q., Ni, H., Xu, H. (2023) Nowcasting Chinese GDP in a data-rich environment: Lessons from machine learning algorithms, *Economic Modelling*, Vol. 122, <https://doi.org/10.1016/j.econmod.2023.106204>.

Приложение 1

Перечень используемых макроэкономических показателей Армении

| Наименование показателя | Источник | Трансформация | Учет сезонности |
|---|----------|---------------|-----------------|
| Объем промышленной продукции, в среднегодовых ценах 2017 г. | Армстат | ln, Δ | Да |
| Объем строительства, в среднегодовых ценах 2017 г. | Армстат | ln, Δ | Да |
| Оборот розничной торговли, в среднегодовых ценах 2017 г. | Армстат | ln, Δ | Да |
| Экспорт товаров | Армстат | ln, Δ | Да |
| Импорт товаров | Армстат | ln, Δ | Да |
| Индекс цен производителей промышленной продукции, % (2010 = 100) | Армстат | ln, Δ | Да |
| Индекс цен строительства, % (2010 = 100) | Армстат | ln, Δ | Да |
| Индекс потребительских цен, % (2010 = 100) | ЦБ РА | ln, Δ | Да |
| Обменный курс AMD/USD, в среднем за период | ЦБ РА | ln, Δ | Да |
| Денежная база (остатки на конец периода) | ЦБ РА | ln, Δ | Да |
| Наличность вне ЦБ РА (остатки на конец периода) | ЦБ РА | ln, Δ | Да |
| Наличность в обращении | ЦБ РА | ln, Δ | Да |
| Денежный агрегат М1 | ЦБ РА | ln, Δ | Да |
| Денежный агрегат М2 | ЦБ РА | ln, Δ | Да |
| Денежный агрегат М2Х | ЦБ РА | ln, Δ | Да |
| Кредиты коммерческих банков резидентам, всего (на конец периода) | ЦБ РА | ln, Δ | Да |
| Кредиты коммерческих банков предприятиям-резидентам, всего (на конец периода) | ЦБ РА | ln, Δ | Да |
| Кредиты коммерческих банков домашним хозяйствам, всего (на конец периода) | ЦБ РА | ln, Δ | Да |
| Средняя ставка по депозитам (сроком до 1 года), в национальной валюте | ЦБ РА | Δ | Нет |
| Средняя ставка по депозитам (сроком до 1 года), в иностранной валюте | ЦБ РА | Δ | Нет |
| Средняя ставка по кредитам (сроком до 1 года), в национальной валюте | ЦБ РА | Δ | Нет |
| Средняя ставка по кредитам (сроком до 1 года), в иностранной валюте | ЦБ РА | Δ | Нет |

Примание: ln — логарифмирование показателя, Δ — взятие первых разностей.

Приложение 2

Перечень используемых макроэкономических показателей Беларуси

| Наименование показателя | Источник | Трансформация | Учет сезонности |
|--|----------|---------------|-----------------|
| Объем промышленного производства, в среднегодовых ценах 2018 г. | НБРБ | ln, Δ | Да |
| Объем инвестиций в основной капитал, в среднегодовых ценах 2018 г. | НБРБ | ln, Δ | Да |
| Объем розничного товарооборота, в среднегодовых ценах 2018 г. | НБРБ | ln, Δ | Да |
| Индекс потребительских цен, в % к предыдущему месяцу | Белстат | ln, Δ | Да |
| Индекс цен производителей промышленной продукции, в % к предыдущему месяцу | Белстат | ln, Δ | Нет |
| Индекс цен производителей сельскохозяйственной продукции, в % к предыдущему месяцу | Белстат | ln, Δ | Да |
| Индекс цен в строительстве, в % к предыдущему месяцу | Белстат | ln, Δ | Да |
| Экспорт товаров и услуг | Белстат | ln, Δ | Да |
| Импорт товаров и услуг | Белстат | ln, Δ | Да |
| Валовые международные резервы | НБРБ | ln, Δ | Нет |
| Номинальный курс белорусского рубля за 1 доллар США | НБРБ | ln, Δ | Нет |
| Денежная база (среднее за месяц) | НБРБ | ln, Δ | Да |
| Широкая денежная масса (среднее за месяц) | НБРБ | ln, Δ | Да |
| Дозиты государственных коммерческих предприятий | НБРБ | ln, Δ | Да |
| Депозиты частного сектора | НБРБ | ln, Δ | Да |
| Депозиты физических лиц | НБРБ | ln, Δ | Да |
| Кредитные вложения банков в государственную собственность | НБРБ | ln, Δ | Да |
| Кредитные вложения банков в частную собственность | НБРБ | ln, Δ | Да |
| Средняя ставка по вновь привлеченным депозитам, в национальной валюте | НБРБ | Δ | Нет |
| Средняя ставка по вновь выданным кредитам (без МБК), в национальной валюте | НБРБ | Δ | Нет |

Примечание: ln — логарифмирование показателя, Δ — взятие первых разностей.



Доклад по суверенному финансированию в 2024 году
(RU/EN)

Суверенное финансирование в Евразийском регионе: рекордные объемы поддержки доноров

Доклад опирается на базу данных, которую ведет ЕФСР, и посвящен мониторингу суверенного финансирования в Евразийском регионе за 2024 год.



Рабочий документ РД/24/5
(RU/EN)

Алгоритмы машинного обучения для краткосрочного прогнозирования темпов прироста реального ВВП

Целью исследования стала оценка точности краткосрочных прогнозов, полученных при помощи методов и алгоритмов машинного обучения, по сравнению с результатами традиционных эконометрических моделей прогнозирования.



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

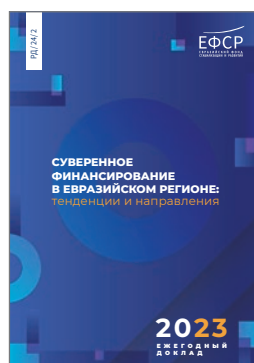
В документе представлены методология и пошаговый алгоритм разработки инструментов для идентификации дисбалансов (кризисов) и стрессовых ситуаций в экономике. Основной акцент сделан на своевременном выявлении нарастания напряженности на валютном рынке.



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

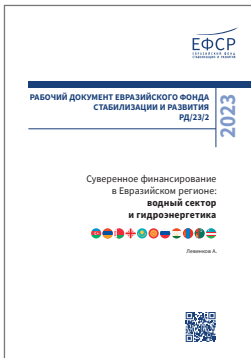
В документе представлен алгоритм действий, который может быть использован для оценки влияния стихийных бедствий на макроэкономические параметры и долговую устойчивость государств.



Рабочий документ РД/23/3
(RU/EN)

Резервы как главный элемент ГСФБ для развивающихся экономик

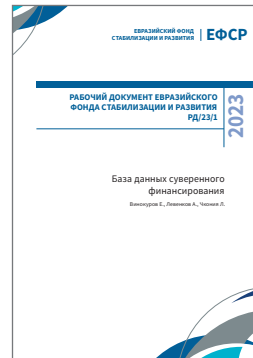
В документе оценены факторы, влияющие на принятие решения развивающимися экономиками об источнике антикризисной поддержки. Анализ показал, что международные резервы являются наиболее востребованным инструментом из всех элементов ГСФБ.



Рабочий документ РД/23/2
(RU/EN)

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

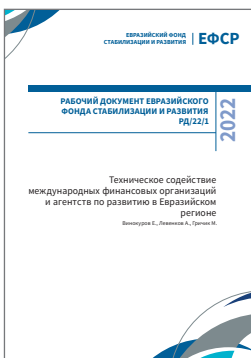
В данном Рабочем документе ЕФСР проведен количественный и качественный анализ операций суверенного финансирования в водно-энергетическом секторе (ВЭС) в 11 странах региона с 2008 по первое полугодие 2023 года.



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.



Наукастинг ВВП: от традиционных эконометрических
моделей к машинному обучению

Цукарев Т., Погосян К., Лемба К., Гришин Д.

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 послемедствий 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

Контакты

Москва,
Чистопрудный бульвар, 17, стр. 1,
101000, Российская Федерация
Тел.: +7 495 775 77 55
Веб: efsd.org



www.efsd.org