

Цифровая экономика

ТЕРРИТОРИАЛЬНОЕ ПЛАНИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ

Ю. А. ПЛЕСКАЧЕВ
Ю. Ю. ПОНОМАРЕВ
М. А. САПРЫКИН

В современных условиях важным элементом планирования территориального и городского развития является своевременный учет прогнозов развития как макро-, так и микроэкономической ситуации в регионе и мегаполисе. В статье предложен подход к прогнозированию экономических показателей, который позволял бы одновременно учитывать динамику макроэкономических факторов и реализацию мероприятий отдельных программных и стратегических документов (на примере мероприятий национальных проектов). При этом проанализированы и показаны наиболее эффективные с точки зрения точности построенных прогнозов варианты построения модельного комплекса из нескольких вариантов современных модельных архитектур на основе их апробации на двух важных для сферы территориального планирования индикаторах — объемах инвестиций в основной капитал и реальных располагаемых доходах населения.

Статья подготовлена в рамках выполнения научно-исследовательской работы государственного задания РАНХиГС при Президенте Российской Федерации.

Ключевые слова: прогнозирование, планирование, машинное обучение.

JEL: C40, C45, R53, E22, D31.

Введение

Развитие городов и городское планирование тесно связаны с учетом актуальной макро- и микроэкономической динамики показателей деятельности экономики, факторов, оказывающих влияние на ее состояние. Так, например, изменения показателя инвестиций в основной капитал отражают возможность предприятий развивать производственные мощности, улучшать инфраструктуру. Значительные колебания и возможные в будущем изменения могут ощутимым образом влиять на принимаемые органами власти и предприятиями решения.

Еще одним важным аспектом, который, помимо изменчивости общих макроэкономических факторов, требует учета для повышения точности анализа и прогнозирования экономических показателей, являются эффекты от реализуемых стратегических и программных

документов, мероприятий в их рамках — например, государственных программ, стратегий развития отраслей, национальных проектов и иных, в рамках которых реализуются масштабные усилия по достижению поставленных целей, в том числе по ключевым макро- и отраслевым показателям. В настоящее время в исследовательской литературе и прикладных аналитических работах этому вопросу, на наш взгляд, уделяется недостаточное внимание.

В данной статье авторами предложен подход к прогнозированию экономических показателей, который одновременно учитывал бы динамику ключевых макроэкономических факторов, неявные и нелинейные взаимосвязи между ними, взаимное влияние и, одновременно, реализацию мероприятий программных и стратегических документов (на

Плескачев Юрий Андреевич, старший научный сотрудник РАНХиГС при Президенте Российской Федерации (Москва), e-mail: pleskachuev-ya@ganepa.ru; Пономарев Юрий Юрьевич, заведующий лабораторией инфраструктурных и пространственных исследований РАНХиГС при Президенте Российской Федерации, канд. экон. наук (Москва), e-mail: ronomarev@ganepa.ru; Сапрыкин Матвей Алексеевич, младший научный сотрудник РАНХиГС при Президенте Российской Федерации (Москва), e-mail: saprykin-ma@ganepa.ru

примере мероприятий национальных проектов). Исходя из общей логики предложенного подхода проведено сравнение относительной точности оценок для нескольких вариантов современных модельных архитектур на основе их апробации на двух важных для сферы территориального планирования индикаторах – объемах инвестиций в основной капитал и реальных располагаемых доходах населения.

Обзор литературы

Вопросы прогнозирования различных макро- и микроэкономических индикаторов широко освещены в исследовательской литературе. Однако лишь некоторые модельные подходы к прогнозированию (включая возможность прогнозирования нескольких показателей) способны учитывать не только набор объясняющих факторов, но и результаты реализации программных и стратегических документов. Далее кратко рассмотрены основные типы моделей, используемых для построения прогнозов и при этом потенциально имеющих возможность учитывать указанный выше аспект, представленных в виде двух больших групп – линейных и нелинейных моделей.

Среди линейных моделей можно выделить линейные регрессионные модели с регуляризацией (Ridge, Lasso) [22]. Из основных целей их применения можно отметить следующие: сокращение размерности регрессоров «в ширину» – с определением таким образом только действительно значимых объясняющих переменных, а также улучшение генерализирующей способности модели [12]. Факторные модели строятся на основе теоретических механизмов, сохраняют интерпретируемость операций с данными и предоставляют возможность количественно оценить влияние отдельного компонента на динамику заданного индикатора [13].

Среди нелинейных моделей можно выделить следующие: многослойный перцептрон [17], сверточную нейронную сеть [15] и градиентный бустинг [8]. Методы машинного обучения активно набирают популярность в эко-

номической сфере для прогнозирования как отдельных показателей, так и одновременно прогноза сразу нескольких заданных индикаторов. Далее подробнее рассмотрен класс нелинейных моделей.

Многослойный перцептрон – полносвязная нейронная сеть, состоящая из нескольких слоев, каждый из которых, в свою очередь, состоит из нейронов нескольких типов: сенсорных – входных, ассоциативных – обучаемых и реагирующих – выходных. В модели присутствует несколько слоев из обучаемых нейронов для решения экономических задач; их количество обычно не превышает двух-трех слоев, в каждом из которых нейронов, как правило, меньше, чем в предыдущем. Такая структура позволяет аппроксимировать нелинейные функции в истинной модели. Нейрон получает на вход взвешенные сигналы, которые суммирует и передает по сети дальше на выход или другому нейрону.

Сверточная нейронная сеть – архитектура, направленная на эффективное распознавание образов, поэтому чаще она используется для решения задач по распознаванию объектов на изображении, однако может применяться не только для решения задач классификации.

Градиентный бустинг – способ объединения (ансамблирования) слабых алгоритмов в один сильный. В рамках проведенного анализа под градиентным бустингом понимается градиентный бустинг над решающими деревьями (*Gradient Boosting on Decision Trees*).

Такие алгоритмы довольно неплохо справляются с гетерогенными данными (к которым можно отнести табличные), что означает поиск нелинейных связей для разнородных признаков (ввиду большой размерности данных в задаче и небольшого количества исторических записей выборку для обучения/тестирования можно назвать таковой). Основная идея подобного метода заключается в том, что на каждом шаге, или итерации, алгоритм все с большей точностью приближается к истинному значению, предсказывая значение ошибки в конце каждой из итераций [5].

Важно отметить, что, несмотря на преимущества и недостатки различных моделей, в большей части исследований не учитывается одновременное влияние макро- или микро-экономических факторов, следующих из структурных моделей, и эффектов от реализуемых стратегических и программных документов, а также мероприятий в их рамках.

Эмпирический подход

Особенностью решаемой задачи является то, что, помимо набора макроэкономических факторов в рамках модельного комплекса, используемого для прогнозирования, учитывались и результаты реализации стратегических и программных документов, в частности мероприятия и параметры национальных проектов, направленных на достижение стратегических целей России.

Территориальное и городское планирование играет важную роль в создании благоприятной среды для инвестирования в основной капитал. Разработка генеральных планов городов и их реализация позволяют создавать комфортные условия для развития бизнеса, повышения конкурентоспособности регионов и привлечения в них новых инвестиций. Активно развивающиеся территории привлекают больше инвестиций в основной капитал, способствуя их экономическому росту и улучшению качества жизни населения.

Высокий уровень реальных располагаемых доходов населения может способствовать росту спроса на жилье, инфраструктуру и услуги, что также может позитивным образом влиять на решения по развитию городской инфраструктуры и реализации инвестиционных проектов. В свою очередь, территориальное планирование и инвестиции в основной капитал, нацеленные на улучшение качества жизни населения и создание благоприятных условий для бизнеса, могут приводить к повышению реальных располагаемых доходов граждан. Таким образом, важно учитывать взаимосвязь рассмотренных факторов.

В связи с этим апробация логики подхода к прогнозированию была проведена нами для показателей инвестиций в основной капитал и реальных располагаемых доходов населения, которые были рассмотрены как целевые (зависимые) переменные в модели (каждая отдельно и совместно в рамках общей спецификации).

Набор объясняющих факторов можно разделить на две группы: экономические факторы, выделенные на основе анализа актуальной исследовательской литературы (представлены в таблице), и факторы, связанные с реализацией национальных проектов – как в части достижения целевых показателей, так и в части расходов на реализацию мероприятий федеральных проектов. Обе группы показателей рассматривались за период с 2008 по 2020 гг.

Для учета влияния мероприятий и параметров национальных проектов на динамику выбранных для прогнозирования индикаторов использовались таблицы расходов Федерального казначейства России [4], в соответствии с которыми на основании целевых статей расходов (и достаточной детализации, включая разделы и подразделы бюджета) определялось, какое мероприятие национального проекта (или иных документов) оказывает влияние на выбранные индикаторы. Благодаря сопоставительным таблицам Минфина России стало возможным найти соответствие расходам на реализацию нацпроектов задолго до их утверждения – до 2012 г.

При охарактеризованном подходе задача исследования состоит в том, чтобы оценить (с помощью моделей регрессии) параметры, описывающие связь между, с одной стороны, объемом расходов конкретной группы, а с другой – показателем, которым описывается достижение национальных целей.

Для оценивания влияния на динамику выбранных индикаторов были рассмотрены следующие варианты архитектуры модельного комплекса, которые можно сгруппировать по критерию размерности вектора зависимой переменной (количество одновременно предсказываемых показателей), – варианты с одной

и двумя переменными, однако архитектура моделей позволяет расширять вторую группу до любого числа заданных переменных.

1. Модели с единичной зависимой переменной (в качестве зависимой переменной рассматривается каждый из описанных выше показателей в отдельности):

- 1.1. Ridge- и Lasso-регрессии.
- 1.2. Catboost-реализация градиентного бустинга на деревьях (CBS).
- 1.3. Многослойный перцептрон (NNS).
- 1.4. Сверточная нейронная сеть (CNNS).
2. Модели с несколькими зависимыми переменными (рассматриваются одновременно

Перечень используемых в моделях факторов для каждой из рассматриваемых зависимых переменных за период 2008–2020 гг.

	Реальные располагаемые денежные доходы населения, в % к предыдущему месяцу	Индекс физического объема инвестиций в основной капитал, в %
Объем работ, выполненных по виду деятельности «Строительство» (в текущих ценах), млн руб.		1
ИПЦ (01.1999=100)	1	1
Цена сырой нефти марки Brent - Европа (средняя за месяц), долл./барр.		1
Ставки по депозитам постоянного действия (средние за месяц), в % годовых	1	
Общая численность занятых в возрасте 15 лет и старше (до 12.2016 – в возрасте 15-72 лет), млн человек		1
Потребность работодателей в работниках, заявленная в органы службы занятости населения, тыс. человек	1	
Индекс глобальной реальной экономической активности		1
Индекс промышленного производства (02.2000=100)		1
Ставка MIACR (1 день) (средняя за месяц), в %		1
Численность граждан, не занятых трудовой деятельностью, состоящих на учете в органах службы занятости населения (по данным Федеральной службы по труду и занятости), тыс. человек	1	
Нагрузка не занятого трудовой деятельностью населения на 100 заявленных вакансий, человек	1	
Индекс продовольственных цен ФАО	1	
Индекс потребительских цен: непродовольственные товары	1	
Платные услуги населению (в текущих ценах)	1	1
Индекс финансового стресса ФРС Сент-Луиса (средний за месяц)	1	
Общая численность безработных в возрасте 15 лет и старше (до 12.2016 – в возрасте 15-72 лет), млн человек	1	
Просроченная задолженность по заработной плате (на начало месяца), млн руб.	1	
Реальная начисленная заработная плата работников организаций (01.1999=100)	1	
Ключевая ставка, в % годовых	1	
Среднедушевые денежные доходы населения, руб.		1
Среднемесячная номинальная начисленная заработная плата работников здравоохранения в организациях, руб. (Росстат)	1	
Средний размер назначенных пенсий по категориям пенсионеров по кварталам, руб.	1	

Источник: составлено авторами по данным КСР ОЭСР.

оба описанные выше показателя — инвестиции в основной капитал и реальные располагаемые доходы населения):

2.1. Catboost-реализация градиентного бустинга на деревьях (CBM).

2.2. Многослойный перцептрон (NNM).

2.3. Сверточная нейронная сеть (CNNM).

Важным отличием моделей с единичной зависимой переменной является то, что факторы из базы данных для каждого рассматриваемого индикатора в случае одновременного прогнозирования нескольких были подобраны индивидуально для каждой из зависимых переменных. Однако для моделей с несколькими зависимыми переменными возможно целиком использовать матрицу ковариат, в которую входят факторы, которые могут оказывать влияние на каждую из зависимых переменных. Это позволяет учесть как возможное взаимное влияние между зависимыми переменными, так и перекрестные связи между независимыми факторами и возможное нелинейное влияние.

Выбор конкретной реализации архитектуры и реализации модельного комплекса из нескольких вариантов производился с учетом следующих требований к подходу:

- высокая точность прогнозов;
- возможность учесть и оценить эффекты от проведения мероприятий в рамках национальных проектов.

Процедура обучения, подбора гиперпараметров и валидации для всех моделей была построена по следующей схеме:

1. Разделение выборки на три части расширяющимся окном: отложенная выборка (последние доступные 12 месяцев — 2021 г.), тренировочная выборка (все доступные данные до тестовой даты) и последние два месяца, доступные в тренировочной выборке, — под подбор гиперпараметров. Такое разделение обусловлено ограничениями в данных (короткий временной ряд, ограниченность и «шоковый характер» данных по группам государственных расходов).

2. Гиперпараметры для моделей подбирались на первой итерации — на этапе, когда

предсказывается первое значение в тестовой выборке.

3. Далее итеративно модель обучалась и делала предсказание на один период вперед, затем по полученным таким образом результатам считались метрики качества ошибок — MAPE и MSE.

В зависимости от архитектуры модели выбирались способы для отбора оптимального сочетания гиперпараметров. Для нейронных сетей использовался один из передовых алгоритмов оптимизации гиперпараметров нейронной сети HyperBand [16], основанный на расширении последовательных алгоритмов двукратного сокращения множества возможных конфигураций гиперпараметров, представленный в платформе подбора гиперпараметров KerasTuner, а для остальных моделей оптимизация гиперпараметров производилась с помощью поиска по сетке.

Полученные результаты

Для построения моделей динамики инвестиций в основной капитал и реальных располагаемых доходов населения использовались несколько спецификаций модельных комплексов.

По результатам сравнительного анализа качества построенных моделей и полученных результатов можно сделать следующие выводы:

- индивидуальный подход к отбору факторов для каждого из показателей дает меньшее значение ошибки прогноза;
- модели на основе градиентного бустинга (CBM/CBS) показали лучший результат по сравнению с другими архитектурными решениями благодаря своим свойствам и особенностям данных (разреженность и большая размерность), работа с которыми для алгоритмов, не основанных на решающих деревьях, вызывает сложности.

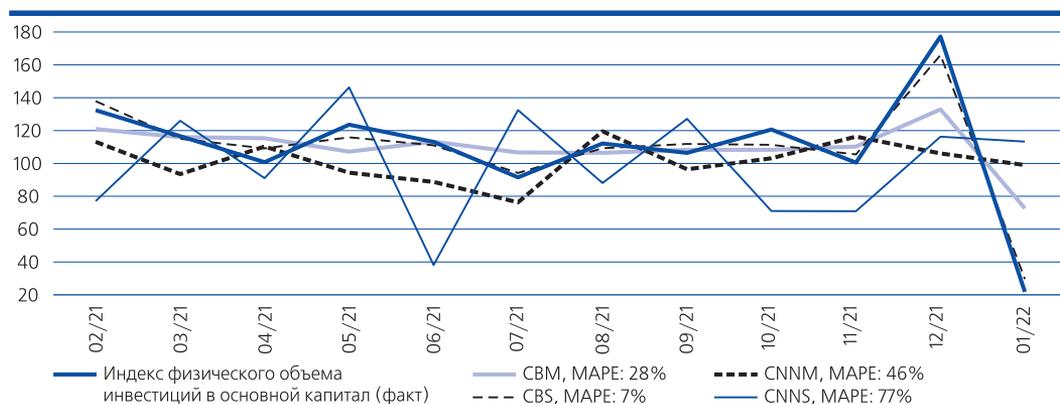
Для более наглядного представления можно рассмотреть соответствующие визуализации результатов работы — прогнозов наиболее точных моделей на тестовой выборке за 2021 г. (См. рис. 1 и 2.)

На рис. 1 можно видеть, что некоторые модели показывают достаточно неплохие результаты, однако при воздействии внешних шоков и высокой чувствительности зависимых переменных к этим шокам прогнозы несколько теряют в своей точности для других специфика-

ций. Так, например, средняя абсолютная процентная ошибка для модели CBS (индивидуальная спецификация для рассматриваемого показателя) на тестовой выборке составила 7%.

Достаточно высокая точность на тестовой выборке модели градиентного бустинга вид-

Рис. 1. Результат работы лучших моделей на тестовой выборке для показателя индекса физического объема инвестиций в основной капитал, в %



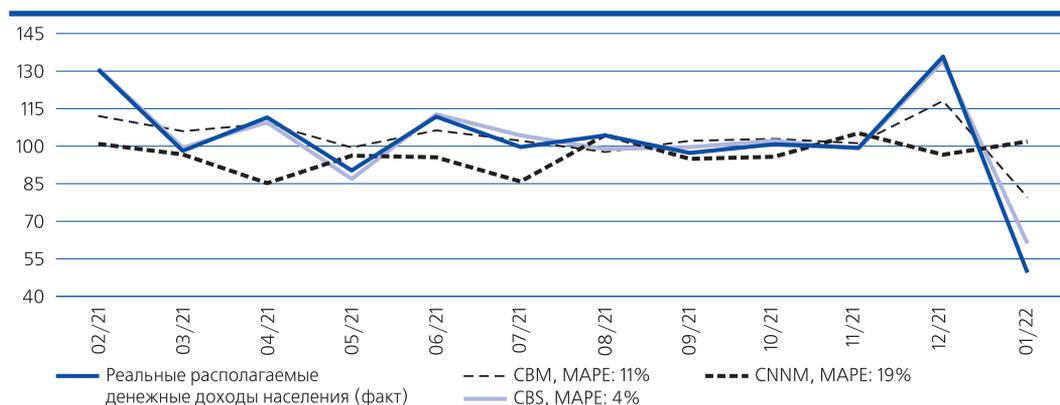
Примечания:

1. CBS, CBM – модели на основе градиентного бустинга (для одной или двух зависимых переменных соответственно); CNNS, CNNM – сверточная нейронная сеть (для одной или двух зависимых переменных соответственно).

2. MAPE – средняя абсолютная процентная ошибка.

Источник: расчеты авторов.

Рис. 2. Результат работы лучших моделей на тестовой выборке для показателя реальных располагаемых денежных доходов населения, в %



Примечания:

1. CBS, CBM – модели на основе градиентного бустинга (для одной или двух зависимых переменных соответственно); CNNM – сверточная нейронная сеть (для двух зависимых переменных).

2. MAPE – средняя абсолютная процентная ошибка.

Источник: расчеты авторов.

на и для второго рассматриваемого показателя — реальных располагаемых денежных доходов населения (см. рис. 2): средняя абсолютная процентная ошибка для модели CBS составила здесь около 4%, для других спецификаций результаты несколько менее обнадеживающие — ошибка составила более 10%.

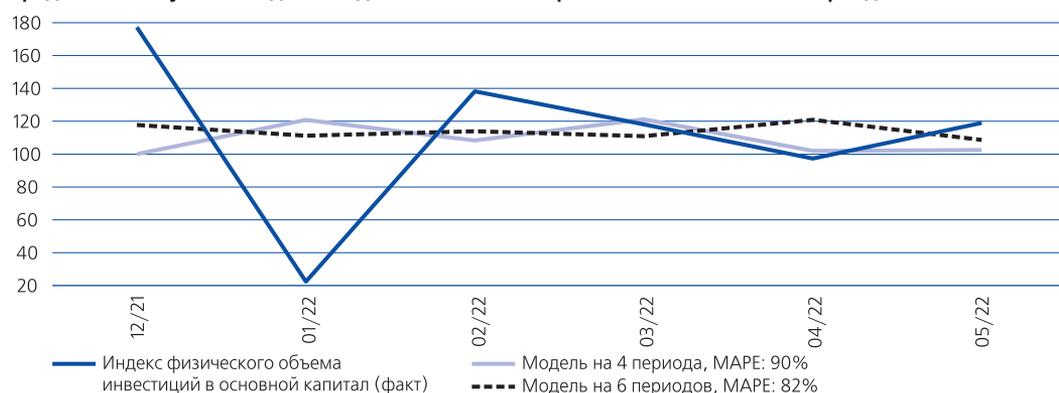
Еще одним элементом анализа стало построение прогнозов вне тестовой выборки — на период 2022 г. В этом случае внешние шоки, которые воздействовали на экономику в 2021-

и, в особенности, в первой половине 2022 г., привели к повышению волатильности рассматриваемых показателей и в результате достаточно негативно сказались на качестве вневыборочных прогнозов для будущих периодов обоих показателей. На рис. 3 и 4 можно видеть, что значения метрик точности достаточно невысоки.

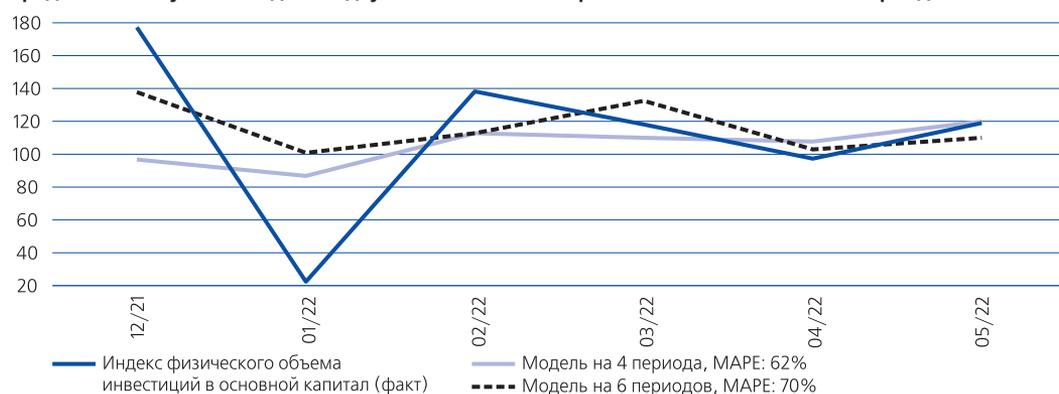
Повышение точности вневыборочных прогнозов возможно после накопления информации в базе данных для обеспечения возможности учета произошедших структурных изме-

Рис. 3. Сравнение вневыборочных прогнозов модели на основе градиентного бустинга для показателя инвестиций в основной капитал на 2022 г., в %

Предсказания лучшей модели с одной зависимой переменной CBS на 4 и 6 периодов



Предсказания лучшей модели с двумя зависимыми переменными СВМ на 4 и 6 периодов



Примечания:

1. CBS, СВМ — модели на основе градиентного бустинга (для одной или двух зависимых переменных соответственно).
2. MAPE — средняя абсолютная процентная ошибка.

Источник: расчеты авторов.

нений во внешних условиях, экономической и политической обстановке.

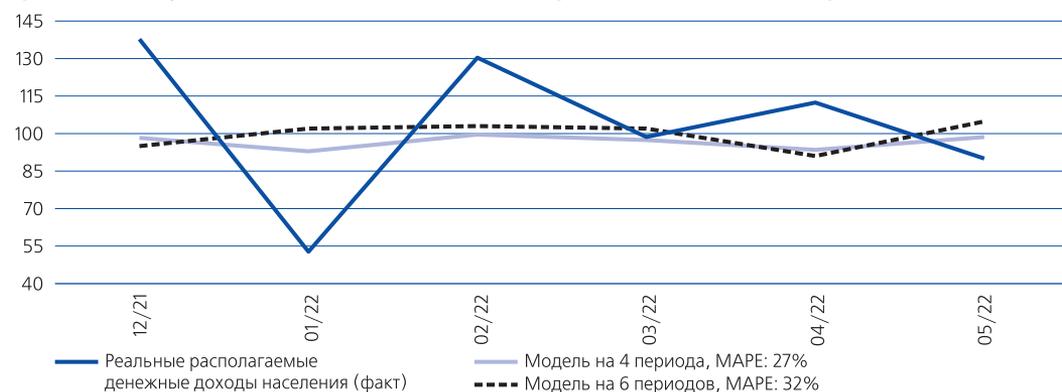
Проведенный анализ в основном сфокусирован на построении кратко- и среднесрочных прогнозов. При дальнейшем же увеличении горизонта прогнозирования значительно увеличивается ошибка прогноза, накапливаемая соответственно за все увеличивающийся период, из-за влияния различных несистемных факторов, которое практически невозможно учесть при построении моделей. Это довольно удачно иллюстрирует сопоставление дина-

мики прогнозов темпов роста ВВП, построенных МВФ в разные годы, с его фактической динамикой. (См. рис. 5.)

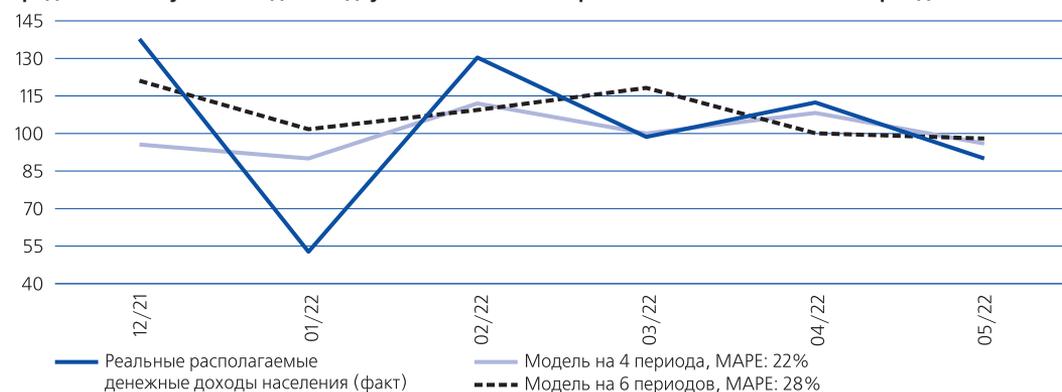
Важно отметить, что одной из ключевых проблем, которая, по нашему мнению, повлияла на качество оценки моделей и полученных на их основе прогнозов, является ограниченность доступных данных, особенно в части параметров и показателей реализации мероприятий в рамках стратегических и программных документов. Одним из решений проблемы с недостатком данных могло бы

Рис. 4. Сравнение вневыборочных прогнозов модели на основе градиентного бустинга для показателя реальных располагаемых денежных доходов населения на 2022 г., в %

Предсказания лучшей модели с одной зависимой переменной CBS на 4 и 6 периодов



Предсказания лучшей модели с двумя зависимыми переменными СВМ на 4 и 6 периодов



Примечания:

1. CBS, SVM — модели на основе градиентного бустинга (для одной или двух зависимых переменных соответственно).

2. MAPE — средняя абсолютная процентная ошибка.

Источник: расчеты авторов.

быть их синтезирование с помощью моделей глубокого обучения генеративно-сопоставительных сетей¹. Такая архитектура представляет собой набор двух нейронных сетей, в которой одна нейросеть создает образцы, похожие на исходные данные, из шума, а вторая — старается отличить, где настоящий исходный массив данных, а где — созданный искусственно. Однако в среднем проведенное оценивание показало, что метрики точности моделей, обученных на данных, полученных при помощи синтетических методов, в итоге ухудшились на 1% по величине ошибки MAPE. Вместе с тем было замечено, что на менее волатильных и подверженных внешним шокам индикаторах ситуация несколько лучше.

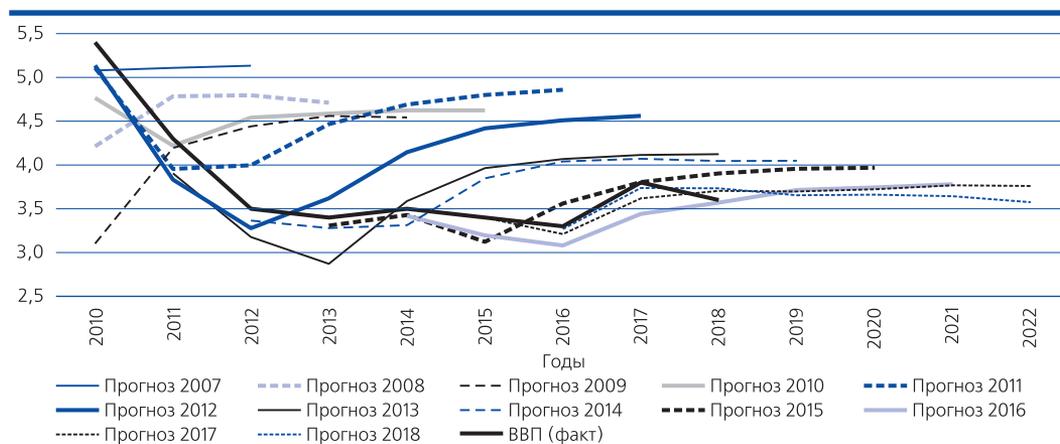
На основе полученных оценок можно заключить, что наличие достаточного количества исторических наблюдений по используемым факторам и большая степень устойчивости к шокам как самих рассматриваемых индикаторов, так и зависимых переменных (факторов) способствовали бы улучшению предсказательных способностей моделей (особенно таких, как градиентный бустинг и нейросетевые архитектуры).

Заключение

В настоящей работе на примере двух показателей, играющих важную роль для территориального и городского планирования и развития — инвестиций в основной капитал и реальных располагаемых доходов населения, рассмотрен механизм построения прогнозов, учитывающих как макроэкономические факторы, так и результаты реализации стратегических и программных документов, с использованием различных методов машинного обучения, для которых проведено сопоставление относительной точности.

Полученные результаты показали, что для работы с разнородной (гетерогенной) матрицей факторов (ковариатов) наилучшим решением в плане точности и эффективности являются модели градиентного бустинга — индивидуальные и мультииндикаторные. Для рассматриваемых модельных комплексов в качестве входящих данных использовалась как информация о макроэкономических факторах, влияющих на соответствующие индикаторы, так и информация о финансировании мероприятий федеральных проектов, с тем чтобы учесть их влияние на такие индикаторы.

Рис. 5. Иллюстрация неточности долгосрочных прогнозов для среднегодовых темпов роста ВВП, в %



Источник: [20].

¹ Generative Adversarial Network (GAN).

Для сопоставления рассмотрено большое количество вариантов моделей и их архитектур, включая модели с единичной зависимой переменной (Ridge- и Lasso-регрессии, реализация градиентного бустинга на деревьях, многослойный перцептрон, сверточная нейронная сеть), а также модели с несколькими зависимыми переменными (реализация градиентного бустинга на деревьях, многослойный перцептрон и сверточная нейронная сеть). Построенные модели были отсортированы по показателю среднего абсолютного процентного отклонения.

Проведя сравнительный анализ качества построенных моделей и полученных результатов, можно сделать следующие выводы:

- индивидуальный подход к отбору факторов позволяет получить наименьшее значение ошибки. В случаях, когда мультииндикаторные модели показывают себя лучше с точки зрения минимизации относительной ошибки, есть ненулевая вероятность того, что модель идентифицировала ложную корреляцию между зависимыми переменными и ковариатами;
- модели на основе градиентного бустинга показали наилучший результат по сравнению с другими архитектурными решениями благодаря своим свойствам и особенностям в данных (разреженность и большая размерность), работа с которыми для алгоритмов, не основанных на решающих деревьях, вызывает сложности;
- с использованием выбранных моделей были построены вневыборочные прогно-

зы для рассмотренных показателей — инвестиций в основной капитал и реальных располагаемых доходов населения на 2022 г. Можно отметить, что с построением прогнозов модели справляются достаточно хорошо, однако при воздействии внешних шоков и высокой чувствительности зависимых переменных к этим шокам прогнозы несколько теряют в своей точности.

Подводя итог, резюмируем, что в настоящей статье авторы, во-первых, на примере выбранных индикаторов представили инструмент прогнозирования, учитывающий не только динамику ключевых макроэкономических показателей, но и результаты выполнения мероприятий программных и стратегических документов (на примере национальных проектов). Полученные выводы могут быть использованы для оперативного анализа и принятия решений, в том числе в сфере территориального планирования.

Во-вторых, представленная методология может использоваться, в том числе, для оценивания эффектов от реализации различных программ развития, поскольку является достаточно гибким инструментом, способным учитывать большое количество разнородных факторов.

В-третьих, архитектура моделей с несколькими переменными позволяет, в случае необходимости, увеличить число прогнозируемых параметров, применяя к ним единую матрицу объясняющих факторов. ■

Литература

1. Беляевский И.К. Денежные доходы населения и потребительские расходы: уровень, тенденции и дифференциация // Статистика и Экономика. 2013. № 2. С. 108–118.
2. Брянцева И.В., Андреева А.И. Прогнозирование инвестиций в основной капитал // Вестник Тихоокеанского гос. ун-та. 2019. Т. 55. № 4. С. 109–118.
3. Ордынская Е.В. Первоочередные задачи государства в обеспечении роста реальных доходов населения // Вестник Российского ун-та дружбы народов. Сер.: Государственное и муниципальное управление. 2019. Т. 6. № 2. С. 75–85.
4. Отчетность по исполнению бюджетов / Федеральное казначейство, 2022. URL: <https://roskazna.gov.ru/ispolnenie-byudzhetov/>
5. Школа анализа данных: учебник по машинному обучению / Сайт Академии Яндекса, 2022. URL: <https://academy.yandex.ru/handbook/ml>

6. Blomstrom M., Lipsey R.E., and Zejan M. Is Fixed Investment the Key to Economic Growth? // *The Quarterly Journal of Economics*. 1996. Vol. 111. No. 1. Pp. 269–276.
7. Bertrand M. et al. Do workfare programs live up to their promises? Experimental evidence from Cote D'Ivoire // *National Bureau of Economic Research*. 2021. No. w28664.
8. Dorogush A.V., Ershov V., Gulin A. CatBoost: gradient boosting with categorical features support // *Электронный архив с открытым доступом для научных статей и препринтов arXiv / Preprint arXiv: 1810.11363*. 2018.
9. Evans W.N., Neville D., and Graham J.D. General Deterrence of Drunk Driving: Evaluation of Recent American Policies // *Risk Analysis*. 1991. Vol. 11. No. 2. Pp. 279–289.
10. Gerdrup K.R., Torstensen K.N. The Effect of Higher Interest Rates on Household Disposable Income and Consumption – A Static Analysis of the Cash-Flow Channel // *Staff Memo Norges Bank*. 2018. No. 3.
11. Herranz Gonzalez F., Martinez-Carrascal C. The impact of firms' financial position on fixed investment and employment: An analysis for Spain // *Banco de Espana. Working Papers*. 2017. No. 1714.
12. Hoerl A.E., Kennard R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems // *Technometrics*. 2000. Vol. 42. No. 1. Pp. 80–86.
13. Huang S.Y., Lee Y.J., Yeh Y.R. Nonlinear dimension reduction with kernel sliced inverse regression // *IEEE transactions on Knowledge and Data Engineering*. 2008. Vol. 21. No. 11. Pp. 1590–1603.
14. Kilinc-Ata N. The evaluation of renewable energy policies across EU countries and US states: An econometric approach // *Energy for Sustainable Development*. 2016. Vol. 31. Pp. 83–90.
15. LeCun Y. et al. Convolutional networks for images, speech, and time series // *The Handbook of Brain Theory and Neural Networks*. 1995. Vol. 3361. No. 10. Pp. 255–258.
16. Li L. et al. Massively parallel hyperparameter tuning // *Электронный архив с открытым доступом для научных статей и препринтов arXiv / Preprint arXiv: 1810.05934*. Vol. 5. 2018.
17. Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain // *Psychol. Rev.* 1958. Vol. 6. No. 65. Pp. 386–408.
18. Sarzynski A., Larrieu J., and Shrimali G. The impact of state financial incentives on market deployment of solar technology // *Energy Policy*. 2012. Vol. 46. Pp. 550–557.
19. Schohn G., Cohn D. Less is more: Active learning with support vector machines // *ICML*. 2000. Vol. 2. No. 4. Pp. 1–6.
20. Seeking sustainable growth. Short-Term Recovery, Long-Term Challenges. *World Economic Outlook (WEO)*. IMF. 2017. 304 p.
21. van der Ploeg F. Disposable income, unemployment, inflation and state spending in a dynamic political-economic model // *Public Choice*. 1989. No. 60. Pp. 211–239.
22. Zhengyan L., Yanbiao X., and Caiya Z. Adaptive Lasso in high-dimensional settings // *Journal of Nonparametric Statistics*. 2009. Vol. 21. No. 10. Pp. 683–696.

References

1. Beliaevskiy I.K. Population income and consumers expenditures: Level, trends and differentiation // *Statistics and Economics*. 2013. No. 2. Pp. 108–118.
2. Bryantseva I.V., Andreeva A.I. Forecasting investments in fixed capital // *Journal of the Pacific State University*. 2019. Vol. 55. No. 4. Pp. 109–118.
3. Ordynskaya E.V. The priority tasks of the state in ensuring the growth of real incomes of the population // *RUDN Journal of Public Administration. Ser.: State and Municipal Administration* 2019. Vol. 6. No. 2. Pp. 75–85.
4. Federal Treasury Reports on the execution of the consolidated budget of the Russian Federation and the budgets of state extra-budgetary funds / *Federal Treasury*. 2022. URL: <https://roskazna.gov.ru/ispolnenie-byudzhetrov/>
5. Machine Learning: Data Analysis Course Materials / *Yandex Academy website*. 2022. URL: <https://academy.yandex.ru/handbook/ml>
6. Blomstrom M., Lipsey R.E., and Zejan M. Is Fixed Investment the Key to Economic Growth? // *The Quarterly Journal of Economics*. 1996. Vol. 111. No. 1. Pp. 269–276.
7. Bertrand M. et al. Do workfare programs live up to their promises? Experimental evidence from Cote D'Ivoire // *National Bureau of Economic Research*. 2021. No. w28664.
8. Dorogush A.V., Ershov V., Gulin A. CatBoost: gradient boosting with categorical features support // *Electronic archive with open access for scientific articles and preprints arXiv:1810.11363*. 2018.

9. Evans W.N., Neville D., and Graham J.D. General Deterrence of Drunk Driving: Evaluation of Recent American Policies // *Risk Analysis*. 1991. Vol. 11. No. 2. Pp. 279–289.
10. Gerdrup K.R., Torstensen K.N. The Effect of Higher Interest Rates on Household Disposable Income and Consumption – A Static Analysis of the Cash-Flow Channel // *Staff Memo Norges Bank*. 2018. No. 3.
11. Herranz Gonzalez F., Martinez-Carrascal C. The impact of firms' financial position on fixed investment and employment: An analysis for Spain // *Banco de España. Working Papers*. 2017. No. 1714.
12. Hoerl A.E., Kennard R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems // *Technometrics*. 2000. Vol. 42. No. 1. Pp. 80–86.
13. Huang S.Y., Lee Y.J., Yeh Y.R. Nonlinear dimension reduction with kernel sliced inverse regression // *IEEE transactions on Knowledge and Data Engineering*. 2008. Vol. 21. No. 11. Pp. 1590–1603.
14. Kilinc-Ata N. The evaluation of renewable energy policies across EU countries and US states: An econometric approach // *Energy for Sustainable Development*. 2016. Vol. 31. Pp. 83–90.
15. LeCun Y. et al. Convolutional networks for images, speech, and time series // *The Handbook of Brain Theory and Neural Networks*. 1995. Vol. 3361. No. 10. Pp. 255–258.
16. Li L. et al. Massively parallel hyperparameter tuning // *Electronic archive with open access for scientific articles and preprints arXiv: 1810.05934*. Vol. 5. 2018.
17. Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain // *Psychol. Rev.* 1958. Vol. 6. No. 65. Pp. 386–408.
18. Sarzynski A., Larrieu J., and Shrimali G. The impact of state financial incentives on market deployment of solar technology // *Energy Policy*. 2012. Vol. 46. Pp. 550–557.
19. Schohn G., Cohn D. Less is more: Active learning with support vector machines // *ICML*. 2000. Vol. 2. No. 4. Pp. 1–6.
20. Seeking sustainable growth. Short-Term Recovery, Long-Term Challenges. *World Economic Outlook (WEO)*. IMF. 2017. 304 p.
21. van der Ploeg F. Disposable income, unemployment, inflation and state spending in a dynamic political-economic model // *Public Choice*. 1989. No. 60. Pp. 211–239.
22. Zhengyan L., Yanbiao X., and Caiya Z. Adaptive Lasso in high-dimensional settings // *Journal of Nonparametric Statistics*. 2009. Vol. 21. No. 10. Pp. 683–696.

Territorial Planning and Forecasting of Economic Indicators by Machine Learning Methods

Yury A. Pleskachyev – Senior Researcher of the Russian Presidential Academy of National Economy and Public Administration (Moscow, Russia). E-mail: pleskachyev-ya@ranepa.ru

Yury Yu. Ponomarev – Head of Laboratory for Infrastructure and Spatial Research of the Russian Presidential Academy of National Economy and Public Administration, Candidate of Economic Sciences (Moscow, Russia). E-mail: ponomarev@ranepa.ru

Matvey A. Saprykin – Junior Researcher of the Russian Presidential Academy of National Economy and Public Administration (Moscow, Russia). E-mail: saprykin-ma@ranepa.ru

Forehanded consideration of economic development forecasts for both macroeconomic and microeconomic situation in the region and the metropolis is an important element in territorial planning and urban development in modern conditions. The article proposes an approach to forecasting economic indicators, which would allow simultaneously taking into account the dynamics of macroeconomic factors and the effects of individual program and strategic documents implementation (using the measures of national projects as an example). Using several options of modern model architectures, we show the most effective model in terms of forecast accuracy based on their approbation on two important indicators for the sphere of territorial planning – investments in fixed assets and real disposable incomes of the population.

The article was prepared as part of the research work of the state task of the RANEPa.

Key words: forecasting, planning, machine learning.

JEL-codes: C40, C45, R53, E22, D31.