



Банк России



ПРОГНОЗИРОВАНИЕ РЕГИОНАЛЬНОЙ ИНФЛЯЦИИ С ПОМОЩЬЮ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ НА ПРИМЕРЕ МАКРОРЕГИОНА СИБИРЬ

№ 91 / март 2022

Серия докладов об экономических исследованиях

О. Семитуркин
А. Шевелев

Оглавление

РЕЗЮМЕ.....	3
1. ВВЕДЕНИЕ.....	4
2. ОБЗОР ЛИТЕРАТУРЫ.....	4
3. ДАННЫЕ.....	6
4. МЕТОДЫ ИССЛЕДОВАНИЯ.....	7
5. РЕЗУЛЬТАТЫ.....	12
6. ЗАКЛЮЧЕНИЕ	16
7. СПИСОК ЛИТЕРАТУРЫ.....	18

Олег Семитуркин

Банк России, Сибирское главное управление

E-mail: SemiturkinON@cbr.ru

Андрей Шевелев

Банк России, Департамент исследований и прогнозирования

E-mail: ShevelevAA@cbr.ru

Авторы выражают признательность Ивану Вербному и анонимным рецензентам за полезные замечания и предложения.

Серия докладов Банка России проходит процедуру анонимного рецензирования со стороны членов Консультативного исследовательского совета Банка России и внешних рецензентов.

Все права защищены. Содержание настоящего доклада отражает личную позицию авторов и может не совпадать с официальной позицией Банка России. Банк России не несет ответственности за содержание доклада. Любое воспроизведение представленных материалов допускается только с разрешения авторов.

Фото на обложке: Shutterstock/FOTODOM

Адрес: 107016, Москва, ул. Неглинная, 12

Телефон: +7 495 771-91-00, +7 495 621-64-65 (факс)

Официальный сайт Банка России: www.cbr.ru

© Центральный банк Российской Федерации, 2022

РЕЗЮМЕ

В работе оценивается качество прогнозирования региональной инфляции при помощи методов машинного обучения на примере макрорегиона Сибирь¹ и сибирских регионов. На первом этапе исследования мы делаем прогнозы региональной инфляции на разные сроки при помощи нескольких методов машинного обучения и методов бенчмарка. На втором этапе мы комбинируем прогнозы методами машинного обучения и взвешиваем их на основе полученных метрик качества. В завершение исследования мы сопоставляем полученные метрики качества с бенчмарками и подтверждаем устойчивость полученных результатов при помощи теста Диболда – Мариано.

По результатам проведенного исследования мы делаем вывод, что качество прогнозирования инфляции в макрорегионе Сибирь и регионах Сибири с помощью методов машинного обучения сопоставимо с традиционными эконометрическими методами. Вместе с тем для каждого региона необходима предварительная оценка качества прогноза методами машинного обучения и целесообразности их использования в сравнении с эконометрическими методами. Для большинства сибирских регионов и для макрорегиона Сибирь на сроках свыше года методы машинного обучения работают лучше бенчмарков, в отличие от прогнозов на один-три квартала вперед. Прогнозирование методом комбинирования моделей машинного обучения в большинстве случаев предпочтительнее, чем каким-то одним.

Ключевые слова: денежно-кредитная политика, машинное обучение, градиентный бустинг, эластичная сеть, прогнозирование временных рядов.

JEL-классификация: E31, E37, E52, E58.

¹ Макрорегион Сибирь – это Сибирский федеральный округ в границах до 2018 г. (то есть Забайкальский край и Республика Бурятия включительно).

1. ВВЕДЕНИЕ

После перехода Банка России к политике таргетирования инфляции точность прогноза темпа роста цен стала необходимым условием для проведения эффективной денежно-кредитной политики. Целью нашей работы является оценка качества прогнозирования инфляции в регионах Сибири с помощью методов машинного обучения.

Современные методы машинного обучения успешно применяются для обработки данных во многих областях, но для прогнозирования макроэкономических показателей (в частности, инфляции) их начали использовать сравнительно недавно. В своем исследовании мы применяем популярные методы машинного обучения, такие как лассо-регрессия, гребневая регрессия, эластичная сеть, случайный лес деревьев решений и градиентный бустинг деревьев решений для прогноза региональной инфляции в регионах Сибири. Для оценки качества прогноза мы сравнили метрики качества прогнозов, полученные методами машинного обучения, с качеством прогноза при помощи широко распространенных для решения подобных задач моделей ARIMA и ARIMAX. Затем получили комбинированный прогноз, построенный на основе взвешенных результатов прогнозирования методами машинного обучения, и провели тест на выявление значимых отличий между методом бенчмарка и методом комбинированного прогноза. По каждой модели качество полученных прогнозов сопоставлено с качеством бенчмарка на основе показателя $(RMSE_x/RMSE_b)$. Множество таких относительных оценок является формальным критерием достижения цели нашего исследования.

2. ОБЗОР ЛИТЕРАТУРЫ

Первые работы с результатами прогнозирования макроэкономических переменных при помощи методов машинного обучения появились сравнительно недавно. С. Chakraborty, A. Joseph (2017) в своей статье приводят различные методы машинного обучения применительно к решению задач центральных банков. В частности, рассматривается пример по прогнозированию инфляции в Великобритании на среднесрочный горизонт – до двух лет. В качестве объясняющих переменных авторы использовали 12 макроэкономических показателей за период с I квартала 1988 г. по IV квартал 2015 г. (112 наблюдений): ВВП, производительность труда, валовой располагаемый доход домохозяйств, денежная масса, долг частного сектора, уровень безработицы, уровень занятости, реальный эффективный обменный курс фунта стерлингов, учетная ставка Банка Англии, разница в доходности индексированных и неиндексированных пятилетних ценных бумаг, глобальный индекс цен на сырьевые товары. В результате авторы приходят к выводу, что методы машинного обучения, такие как нейронная сеть, метод опорных векторов, модель случайного леса и др., показывают качество прогноза выше, чем традиционно используемые для этих целей методы – например, модель авторегрессии AR(1).

J. Jung et al. (2018) использовали три различных метода машинного обучения для прогнозирования роста реального ВВП уже в семи странах (США, Великобритании, Германии, Мексике, Филиппинах, Вьетнаме, Испании). Выборка включала 29 макроэкономических показателей на квартальной и годовой основе с 70–90-х гг. по 2016 год. В работе также сделан вывод, что приведенные методы машинного обучения стабильно превосходят методы, традиционно используемые в прогнозах МВФ.

В работе И. Байбуза (2018) доказываемость состоятельность ML-методов при прогнозировании инфляции в России в сравнении с традиционными альтернативами. Автор использует временной ряд ИПЦ и выборку из 92 макроэкономических ежемесячных показателей с февраля 2002 г. по июнь 2016 г. (173 наблюдения). Выбранные данные отражают состояние деловой активности, промышленного производства, денежного рынка, уровня занятости, платежного баланса, а также цен на основные экспортные товары российской экономики. Методы случайного леса и бустинга при прогнозировании месячной инфляции показали результаты, сопоставимые с базовой моделью AR(1). При этом они показали существенно лучшие результаты при сравнении прогнозов средней инфляции на горизонте более двух месяцев. Автор делает вывод о перспективности упомянутых выше моделей в применении к задаче прогнозирования инфляции в России. Кроме того, автор доказывает, что ансамблевые модели на нетрансформированных данных показывают результаты прогнозирования лучше, чем аналогичные модели, использующие трансформированные данные.

Работа К. Яковлевой (2018) посвящена описанию методики расчета высокочастотного индикатора, отражающего динамику экономической активности в России. В качестве исходных данных используются новостные статьи, взятые из интернет-ресурса, посвященного обзору экономических событий в России и за рубежом. Анализ новостных статей производится с помощью методов текстового анализа и машинного обучения. Сначала автор с помощью вероятностной модели LDA строит индикатор, отражающий количественную составляющую тем, то есть как часто тема упоминается в новостях. Затем автор строит индикатор, который характеризует эмоциональную составляющую новостей, определяя ее с помощью методов машинного обучения с учителем. По полученным индикаторам автор прогнозирует индекс деловой активности. Результаты расчетов показали, что использование новостей может быть важной составляющей при прогнозировании экономической активности в России, о чем свидетельствует хорошее качество полученной модели.

Статья Е. Павлова (2020) посвящена изучению применимости нейронных сетей для прогнозирования инфляции в России. Автор включает в выборку 10 макроэкономических показателей с февраля 2002 г. по август 2018 г. (200 наблюдений): ИПЦ, индекс физического объема ВВП, производительность труда (отношение реального ВВП к числу занятых), агрегат денежной массы M2, объем выданных кредитов в реальном выражении, норма безработицы, экспорт в реальном выражении, цена на нефть в долларах США, реальный располагаемый доход, процентная ставка денежного рынка. Рассмотренные модели показали хорошие результаты на горизонте планирования более одного месяца, что делает их перспективным инструментом для кратко- и среднесрочного прогнозирования. В целом методы машинного обучения, такие как нейронная сеть и метод опорных векторов, показали себя лучше, чем однофакторная и простая линейная модели.

В статье O. Barkan et al. (2021), подготовленной при участии Банка Израиля, авторы оценивают качество прогнозов компонентов инфляции в США при помощи нейронных сетей. Выводом исследования является превосходство качества прогноза нейронными сетями перед AR(p) и VAR(p), особенно на нижних уровнях иерархии индекса потребительских цен США.

Авторы O. Ozgur et al. (2021) в своем исследовании сравнили прогностическую способность уровня инфляции в Турции при помощи методов с регуляризацией с

методами ARIMA и VAR. В результате методы машинного обучения показали более качественный прогноз на среднесрочном периоде по сравнению с бенчмарками.

В докладе Банка России «Прогнозирование ИПЦ в России на винтажных данных и с использованием методов машинного обучения» авторы М. Мамедли и Д. Шибитов показывают, как прогностические способности моделей меняются в зависимости от использования винтажных данных и сезонного сглаживания. Авторы используют метод эластичной сети, ансамблевые методы машинного обучения, нейронные сети и сопоставляют результаты с методом прогнозирования AR(p). Показано, что фактическая ошибка прогноза может быть значительно недооценена при проведении эксперимента в псевдореальном времени. Подчеркнута важность их учета при построении прогнозов в реальном времени.

Отличие нашего исследования от подобных состоит не только в применении методов машинного обучения для прогнозирования региональной инфляции в России, но и в использовании метода комбинирования взвешенных прогнозов моделями машинного обучения и сопоставлении их с результатами, полученными отдельными методами машинного обучения и методами бенчмарка ARIMA и ARIMAX. В качестве факторов у нас используется курс валюты, прошлая динамика региональной инфляции и производные от них (лагированные и средние значения, стандартные отклонения). В аналогичных исследованиях для машинного обучения используются массивы из множества различных факторов, что может значительно осложнить переход от прогнозирования в режиме псевдореального времени к реальному. Кроме того, более широкий набор данных по сравнению с бенчмарками делает непонятным источник улучшения точности прогнозов – метод или данные. В нашем исследовании и в бенчмарках, и в машинном обучении используется один и тот же набор информации, что позволяет более корректно сделать выводы об относительном качестве методов прогнозирования.

3. ДАННЫЕ

В качестве исходных данных мы использовали ИПЦ² за период с февраля 2002 г. по декабрь 2020 г. по макрорегиону Сибирь и 12 сибирским регионам. Ежедневные данные валютного курса преобразованы в ежемесячные путем нахождения среднемесячного значения. Для методов с регуляризацией L1 и L2 исходные ряды предварительно нормировались. Как показал обзор литературы, для ансамблевых методов в аналогичных задачах на практике используются разные подходы к предварительной обработке данных, но предпочтительнее использование нескорректированных данных. Мы также получили лучшие оценки прогнозов ансамблевыми методами на исходных нескорректированных данных по региональной инфляции.

В качестве предикторов для методов машинного обучения использовались ряды ИПЦ и курса национальной валюты в форме лагированных значений от 1 до 12 месяцев, средних значений и стандартных отклонений за предыдущие три, шесть, девять и 12 месяцев. Так, мы получили набор факторов, основанный на двух статистических рядах.

² Федеральная служба государственной статистики. URL: <https://fedstat.ru/> (дата обращения: 12.02.2021).

4. МЕТОДЫ ИССЛЕДОВАНИЯ

В исследовании использовался принцип обучения с учителем. Исходная совокупность данных была разбита нами на тренировочную, валидационную и тестовую выборки. Для решения проблемы переноса информации из тестовой выборки в результат и для подбора гиперпараметров модели использовался поиск оптимальных параметров с использованием кросс-валидации (Рисунок 1).

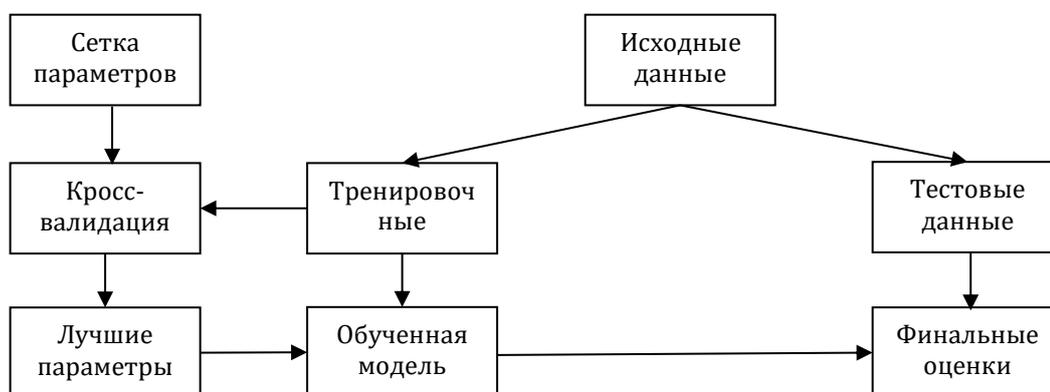
Рис. 1. Принцип кросс-валидации



Источник: составлено авторами.

Таким образом, мы избежали переноса в модель информации из тестового набора и получили более объективные оценки. Изначально ряд разбивался на тренировочную и тестовую компоненты, после чего экспертным способом задавалась сетка гиперпараметров модели. Далее обучающий набор данных разбивался на заданное количество проверочных наборов данных (валидационных), на каждом из которых проводилось обучение модели. По каждому набору гиперпараметров на валидационных выборках вычислялось среднее значение правильности перекрестной проверки. В итоге выбиралась комбинация с лучшими оценками качества. Общий процесс разбиения данных и оценки итоговых параметров представлен на схеме (Рисунок 2).

Рис. 2. Процесс разбиения данных и оценки итоговых параметров



Источник: <https://scikit-learn.org>³.

Прогнозы ИПЦ посчитаны тремя методами с L1-, L2-регуляризацией (лассо-регрессия, гребневая регрессия, эластичная сеть) и двумя ансамблевыми методами (случайный лес

³ На сайте представлено описание библиотеки scikit-learn языка программирования Python.

деревьев решений и градиентный бустинг деревьев решений). Результаты сопоставлялись с помощью двух методов бенчмарка (ARIMA, ARIMAX). Вкратце охарактеризуем специфику использования и практику применения каждого метода прогнозирования.

Для прогноза временного ряда ИПЦ методами ARIMA и ARIMAX лучшие параметры подбирались путем выбора оптимального значения на отложенной выборке (аналогично с методами машинного обучения). Сделано это с целью лучшей сопоставимости прогнозов бенчмарка и машинного обучения. С помощью расширенного теста Дики – Фуллера определен порядок разности для приведения данных к стационарному виду. Длина лагов при выборе гиперпараметров ARIMA, ARIMAX задавалась в диапазоне 1–24. В модели ARIMAX в качестве расширения применен номинальный курс рубля, в остальном принцип выбора лучших параметров модели схож с описанным выше.

Методы, использующие L1- и L2-регуляризацию, являются развитием метода наименьших квадратов. Помимо ограничений на величину квадратичной ошибки, в них добавляется еще одно ограничение, которое в случае гребневой регрессии минимизирует величину коэффициентов при регрессорах, а в лассо-регрессии приравнивает незначительные коэффициенты к нулю. Фактически при помощи лассо-регрессии мы отбираем наиболее важные для модели регрессоры. При этом коэффициенты должны сохранять максимальную прогнозную силу полученной модели. Качество модели оценивается следующим образом:

$$Q = \sum_{i=1}^l L(y_i, a(x_i)) + \gamma V(w) \rightarrow \min_w \quad (1),$$

где Q – показатель качества модели, L – функция потерь, V(w) – регуляризатор модели, y_i – реальные значения функции, γ - параметр настройки веса регуляризатора в модели, $a(x_i)$ – прогнозные значения функции, w – коэффициенты при регрессорах модели, l – размерность векторов параметров модели.

Для гребневой регрессии функция потерь записывается в виде:

$$L(y_i, a(x_i)) = (y_i - a(x_i))^2 \quad (2),$$

регуляризатор L2 в виде:

$$V(w) = \|w\|_{l_2}^2 = \sum_{n=1}^d w_n^2 \quad (3),$$

Для лассо-регрессии функция потерь записывается в виде:

$$L(y_i, a(x_i)) = |y_i - a(x_i)| \quad (4),$$

регуляризатор L1 в виде:

$$V(w) = \|w\|_{l_1} = \sum_{n=1}^d |w_n| \quad (5),$$

Введение регуляризатора позволяет снизить склонность к переобучению и улучшить обобщающие свойства модели по сравнению с линейной. В нашей задаче мы определяем вес регуляризатора путем перебора из заранее заданных значений, оценки и выбора лучших характеристик модели.

Модель эластичной сети (ElasticNet) объединяет свойства лассо-регрессии (Lasso) и

гребневой регрессии (Ridge), то есть использует L1- и L2-регуляризацию с определенными весами. Регуляризатор эластичной сети записывается в виде:

$$V(w) = \vartheta \sum_{n=1}^d |w_n| + (1 - \vartheta) \sum_{n=1}^d (w_n)^2 \quad (6) ,$$

где $\vartheta \in [0; 1]$ – гиперпараметр модели, который определяет вес L1- и L2-регуляризации. В нашей работе мы, как и Chakraborty and Joseph (2017), приняли его равным 0,5.

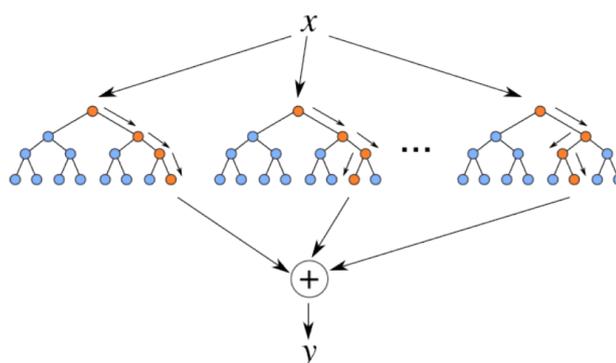
В целом все методы с L1- и L2-регуляризацией отличаются своей простотой, скоростью работы и возможностью интерпретируемости полученных результатов. Особенности применения таких моделей является необходимость настройки параметров регуляризации, устранение возможного переобучения модели. Кроме того, признаки перед использованием в модели необходимо нормировать.

Для ансамблевых методов машинного обучения базовым является метод дерева решений (Decision Tree), который в общем случае представляет из себя граф, состоящий из корневых и листовых вершин. При выборе параметра перехода из одной вершины к другой минимизируется критерий информативности. В случае решения задачи регрессии цель состоит в том, чтобы минимизировать среднеквадратичную ошибку (MSE) в каждой вершине построенного дерева.

$$H(Q_m) = \frac{1}{N_m} \sum_{y \in Q_m} (y - \bar{y}_m)^2 \quad (7) ,$$

где $H(Q_m)$ – критерий информативности (среднеквадратичная ошибка), y – реальные значения функции, \bar{y}_m – полученные оценки функции, N_m – количество значений. Главный недостаток дерева решений – возможность неконтролируемого переобучения и низкая обобщающая способность, поэтому в настоящее время вместо деревьев решений обычно используют случайный лес деревьев решений (Random Forest). Суть этого метода в том, что на выборке строится не одно, а несколько деревьев решений. Для каждого дерева случайным образом выделяется подвыборка данных, на которой строится модель, полученные параметры усредняются. Таким образом, значительно снижается возможность переобучения модели, как на отдельном дереве решений. В отличие от метода дерева решений, в узлах производится случайный отбор признаков, и на их основе происходит поиск оптимального решения (Рисунок 3).

Рис. 3. Схема работы модели случайного леса

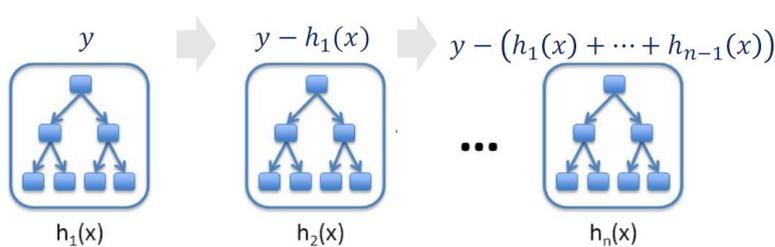


Источник: <https://scikit-learn.org>.

В случае регрессии алгоритм дает прогноз для каждого дерева, после этого полученные значения усредняются. К плюсам данного метода можно отнести устойчивость к переобучению, распределенное обучение модели, отсутствие необходимости предварительной подготовки рядов признаков. Главным минусом является умеренная по сравнению с деревом решений интерпретируемость полученных результатов.

В методе градиентного бустинга деревьев решений (Gradient Boosting Decision Trees) используется множество деревьев решений, но, в отличие от модели случайного леса, они строятся не на подвыборках данных, а на полной выборке. После построения первого дерева мы получаем решения $h_1(x)$, на втором шаге мы строим дерево, в котором в качестве входного параметра используем ошибки первого $y - h_1(x)$. Таким образом, на втором шаге мы ищем решение $h_2(x)$ для минимизирования ошибок первого дерева. Таких шагов мы можем сделать сколько потребуется для получения необходимых параметров модели (Рисунок 4).

Рис. 4 Схема работы модели градиентного бустинга



Источник: <https://scikit-learn.org>.

Основными достоинствами метода являются высокие результаты прогнозирования, отсутствие необходимости предварительной обработки данных, минусы – большая требовательность к точной настройке параметров, плохая работа с разреженными данными, низкая интерпретируемость результатов. Градиентный бустинг является одним из самых мощных и результативных методов прогнозирования, используемых в различных областях машинного обучения. В нашей задаче прогнозирования региональной инфляции он также показал один из лучших результатов.

По полученным данным для каждой модели рассчитана метрика качества прогноза Root Mean Squared Error ($RMSE_x$) и соотнесена с $RMSE_b$ бенчмарка.

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (a_i - y_i)^2} \quad (8),$$

где m – количество точек прогноза, a_i – прогнозное значение ИПЦ, y_i – реальное значение ИПЦ на тестовой отложенной выборке.

Таким образом, для каждой модели получен показатель $RMSE_x/RMSE_b$. Если полученный относительный показатель меньше единицы, $RMSE$ -модели лучше бенчмарка и наоборот.

После подготовки прогнозов упомянутыми методами машинного обучения мы получили комбинированный прогноз. Метод основан на усреднении нескольких прогнозов с учетом качества их прогноза. Вес прогноза каждой модели рассчитывается, исходя из величины обратной оценке $RMSE_x$, отнесенной к сумме аналогичных оценок по всем моделям:

$$W_x = \frac{1/RMSE_x}{\sum_{x=1}^N 1/RMSE_x} \quad (9),$$

$$\sum_{x=1}^N W_x = 1 \quad (10),$$

где W_x – вес прогноза модели x в комбинированном прогнозе; N – количество моделей машинного обучения, в составе комбинированного.

Такой подход позволяет увеличивать вклад лучших моделей с минимальным $RMSE$ и минимизировать вес прогноза при помощи моделей по мере ухудшения качества, показанного на отложенной выборке.

В качестве альтернативной оценки устойчивости полученных результатов мы провели тест Диболда – Мариано (Diebold, Mariano, 1995), чтобы выявить значимые отличия между двумя методами – бенчмарка и комбинированного прогнозирования. Этот тест устойчив к отклонениям от предположений о нормальности, автокорреляции и отличии от нулевого среднего уровня ошибок. Нулевой гипотезой теста является отсутствие различий между прогнозными свойствами двух моделей, то есть равенство нулю среднего уровня разницы между функциями потерь сравниваемых моделей (\bar{d}):

$$\bar{d} = \frac{1}{T} \sum_{t=1}^T (g(e_{At}) - g(e_{Bt})) \quad (11),$$

где $g(e_{At})$ – функция потерь модели А; $g(e_{Bt})$ – функция потерь модели В; T – горизонт прогнозирования временного ряда.

В качестве функции потерь в нашем случае выступает среднеквадратичная ошибка MSE. Решающая статистика (S) теста выглядит следующим образом:

$$S = \frac{\bar{d}}{\sqrt{\hat{\sigma}^2/T}} \quad (12)$$

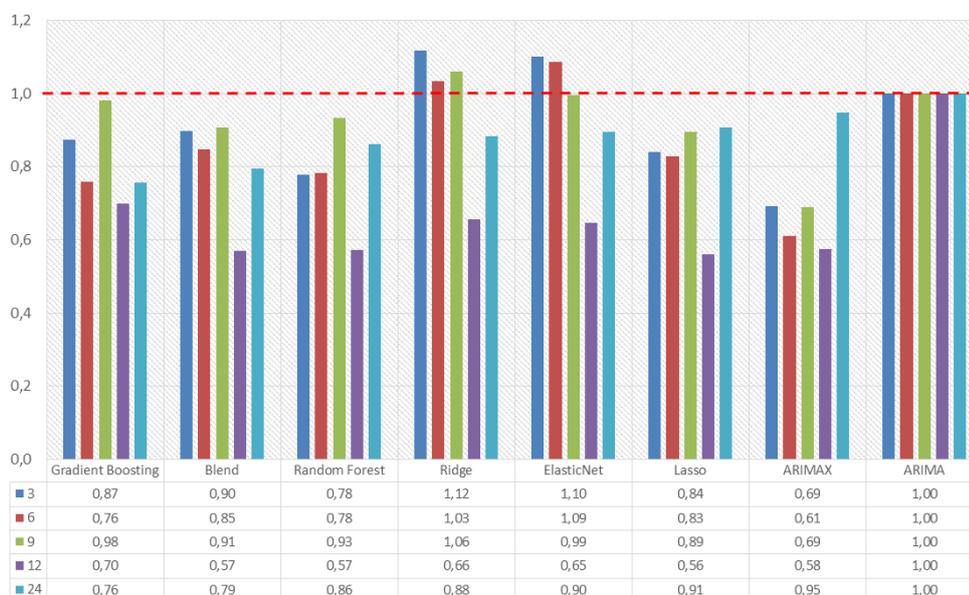
где $\hat{\sigma}^2$ – оценка дисперсии ряда разностей значений функций потерь; T – горизонт прогнозирования временного ряда.

В нашей работе мы оцениваем p-value как вероятность того, что нулевая гипотеза принимается и комбинированный прогноз, полученный на основе моделей машинного обучения, и прогноз на основе модели ARIMAX значимо не различаются.

5. РЕЗУЛЬТАТЫ

Для ИПЦ по макрорегиону Сибирь (Рисунок 5) большинство моделей машинного обучения на горизонте прогнозирования два года показали результаты лучше бенчмарков. Причем общая тенденция сводится к тому, что на сроках год и менее машинное обучение показывает сопоставимые оценки с моделями бенчмарка. Модель комбинированного прогноза и все используемые нами методы машинного обучения показали результаты среднесрочного прогноза на два года лучше бенчмарков.

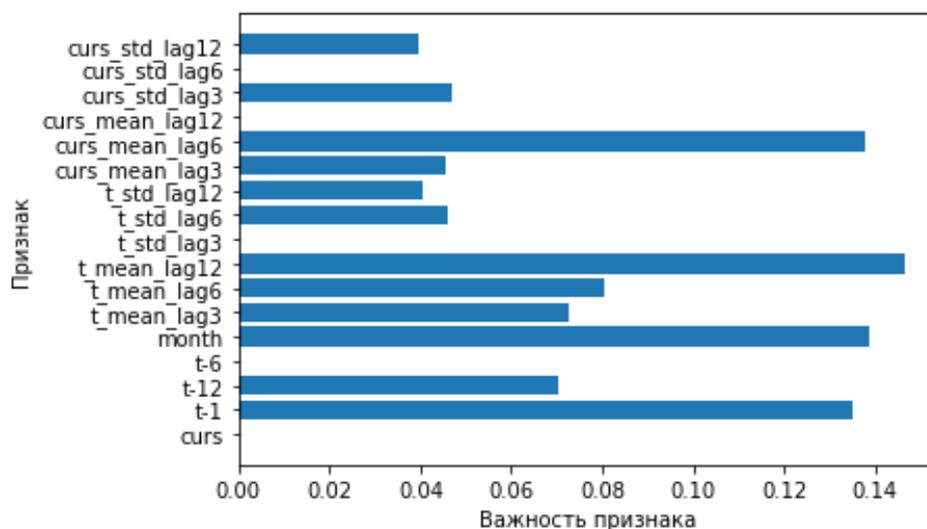
Рис. 5. Оценки качества прогнозов ИПЦ по макрорегиону Сибирь ($RMSE_x/RMSE_b$)



Источник: расчеты авторов.

В группу факторов, наибольшим образом влияющих на прогнозируемый показатель в модели градиентного бустинга (Рисунок 6), попали такие факторы, как среднее значение ИПЦ за последние 12 месяцев (t_mean_lag12), предыдущее значение ИПЦ ($t-1$). Фактор month представляет собой порядковый номер месяца в году и широко используется на практике в подобного рода моделях. Номинальный курс лучшим образом учитывается в модели в виде среднего значения за последние шесть месяцев ($curs_mean_lag6$). Показатели стандартного отклонения по ИПЦ и номинальному курсу вносят наименьший вклад в качество модели.

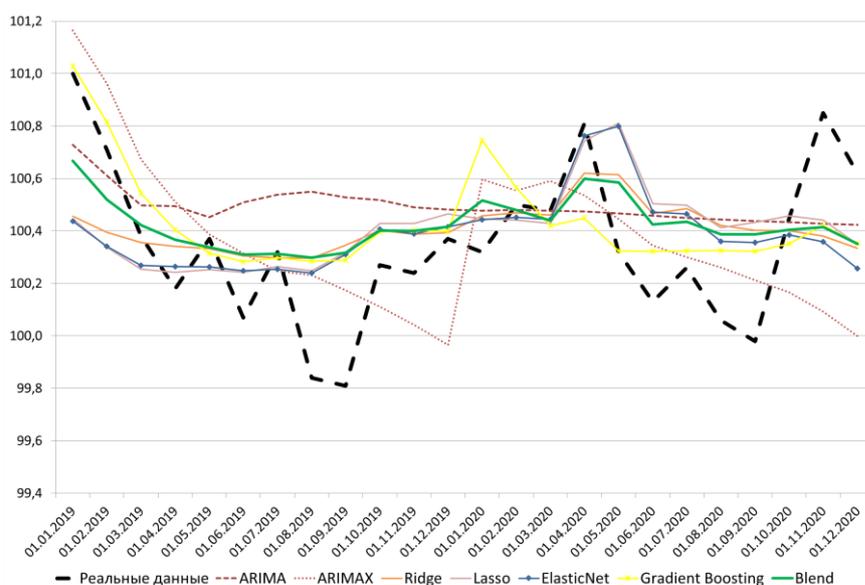
Рис. 6. Значимость факторов модели градиентного бустинга для макрорегиона Сибирь



Источник: расчеты авторов.

Графическое сопоставление реального ряда отложенной выборки и прогнозов, полученных по результатам проведенного исследования данных по макрорегиону Сибирь на 24 месяца (Рисунок 7), также демонстрирует, что прогнозы методами машинного обучения показывают лучшую прогнозную силу. Ускорение инфляции в апреле 2020 г. в наибольшей мере спрогнозировано методами с L1- и L2-регуляризацией эластичной сети (ElasticNet). Модели бенчмарка в основном показывают общее направление динамики инфляции, но какие-то локальные кратковременные ускорения или замедления не находят отражения в прогнозах с помощью этих методов.

Рис. 7. Прогнозирование ИПЦ для макрорегиона Сибирь на 24 месяца



Источник: расчеты авторов.

Усредненные результаты оценки качества моделей по 12 сибирским регионам представлены в Таблице 1, значения в которой отсортированы по столбцу с прогнозами на 24 месяца. В ячейках приведены относительные оценки качества моделей в разрезе сроков

прогнозов и методов моделирования относительно модели ARIMA, качество которой принято за единицу. Чем ниже значение в ячейке, тем меньше $RMSE_x$ модели относительно ARIMA, а значит, лучше прогноз.

Табл. 1. Усредненные результаты по регионам макрорегиона Сибирь

Среднее значение $RMSE_x/RMSE_y$					
Срок прогноза (месяцев)	3	6	9	12	24
Gradient Boosting	0,87	0,76	0,98	0,70	0,76
Blend	0,90	0,85	0,91	0,57	0,79
Random Forest	0,78	0,78	0,93	0,57	0,86
Ridge	1,12	1,03	1,06	0,66	0,88
ElasticNet	1,10	1,09	0,99	0,65	0,90
Lasso	0,84	0,83	0,89	0,56	0,91
ARIMAX	0,69	0,61	0,69	0,58	0,95
ARIMA	1,00	1,00	1,00	1,00	1,00

Источник: расчеты авторов.

На обобщенных данных по сибирским регионам подтвердилась гипотеза об улучшении прогнозов методами машинного обучения на сроках 12 месяцев и 24 месяца. Метод комбинированного прогноза (Blend) на горизонте 12 месяцев и 24 месяца показал лучший результат практически наравне с Lasso и Random Forest.

В качестве дополнительной оценки мы провели тест на отсутствие значимых различий между прогнозами бенчмарка (ARIMAX) и методом комбинированного прогноза (Blend) ИПЦ макрорегиона Сибирь. В качестве нулевой гипотезы мы принимали то, что прогнозы этих моделей значимо не отличаются. По результатам получены следующие значения p_value для нулевой гипотезы (Таблица 2).

Табл. 2. Результаты теста Диболда – Мариано прогнозов ИПЦ макрорегиона Сибирь

Срок прогноза (месяцев)	3	6	9	12	24
P_value DM_test (Blend vs ARIMAX)	0,04	0,1917	0,1856	0,15	0,017

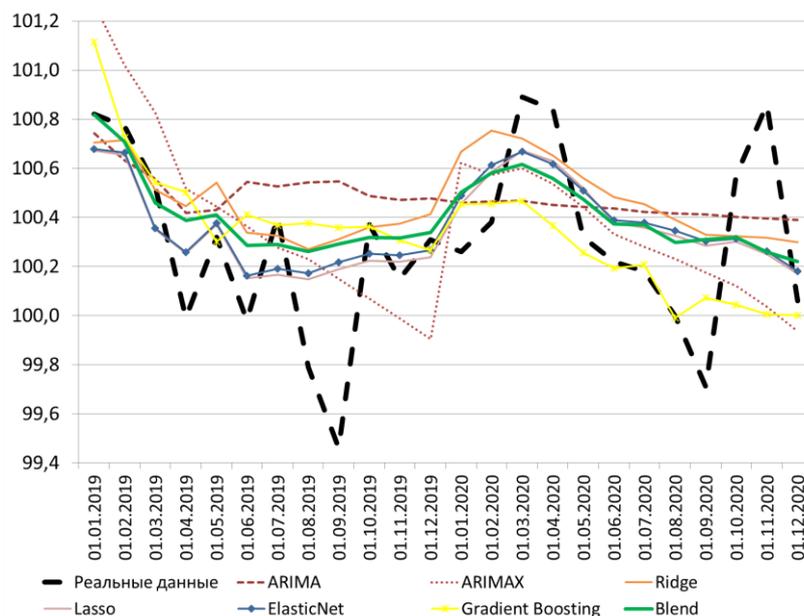
Источник: расчеты авторов.

Результаты позволяют отвергнуть нулевую гипотезу на 95%-ном уровне значимости для прогнозов методами комбинированного прогноза и ARIMAX на три и 24 месяца. Причем на горизонте три месяца оказывается точнее прогноз ARIMAX, а на горизонте 24 месяца – комбинированный прогноз. Для других сроков прогнозирования принять альтернативную гипотезу на 95%-ном уровне значимости мы не можем. Полученный результат подтверждается выводами на основе отношений $RMSE_x/RMSE_y$. Таким образом, можно предположить, что между прогнозами бенчмарка (ARIMAX) и методом комбинированного прогноза (Blend) ИПЦ макрорегиона Сибирь есть значимые различия и метод комбинированного прогноза уместно использовать для решения подобной задачи.

Прогнозирование ИПЦ по 12 регионам макрорегиона Сибирь в среднем также показало лучшие результаты методами машинного обучения на сроках прогнозирования год и более. В качестве примера приведем прогнозы для Новосибирской области на 24 месяца (Рисунок 8). Модель ARIMAX показала близкий по качеству прогноз к методам

машинного обучения. Лучшие результаты показали методы с L1- и L2-регуляризацией и комбинирования моделей машинного обучения за счет более качественного прогноза динамики I квартала 2020 года.

Рис. 8. Прогнозирование ИПЦ для Новосибирской области на 24 месяца



Источник: расчеты авторов.

По Новосибирской области на сроках больше года модели машинного обучения также работают лучше бенчмарков (Таблица 3).

Табл. 3. Оценки качества прогнозов ИПЦ по Новосибирской области

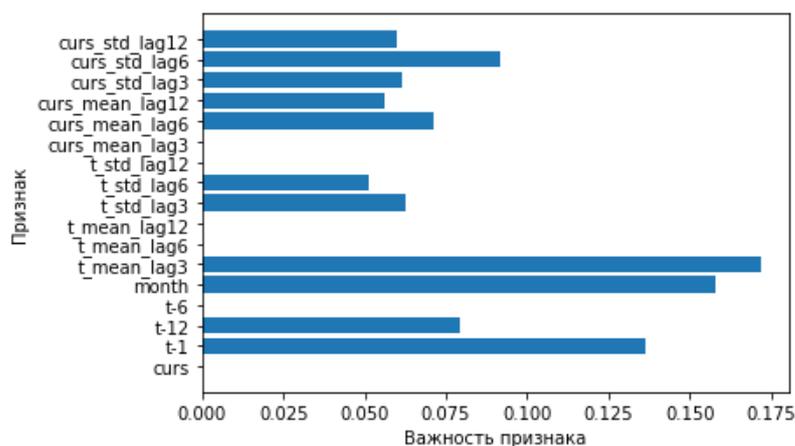
RMSE _x /RMSE _б					
Метод / Срок прогноза (месяцев)	3	6	9	12	24
Blend	0,94	1,05	0,91	0,73	0,87
ElasticNet	0,92	1,03	0,88	0,64	0,89
Gradient Boosting	0,86	1,10	0,96	0,88	0,89
Lasso	0,99	1,06	0,94	0,77	0,90
Ridge	1,07	1,11	1,03	0,80	0,91
ARIMAX	0,70	0,97	1,13	0,72	0,96
ARIMA	1,00	1,00	1,00	1,00	1,00
Random Forest	0,99	1,13	0,96	0,81	1,16

Источник: расчеты авторов.

На периодах один и два года прогнозы, полученные методами машинного обучения, сопоставимы с качеством прогнозов методами бенчмарка. Можно сделать вывод о возможности использования методов машинного обучения для прогноза ИПЦ по Новосибирской области на 24 месяца методами комбинированного прогноза, эластичной сети (ElasticNet) и градиентного бустинга (GradientBoosting) совместно с методом бенчмарка ARIMAX.

Метод градиентного бустинга в большей мере использует в прогнозе усредненные данные и стандартные отклонения валютного курса и исходного ряда ИПЦ (Рисунок 9). Наиболее важными факторами в этой модели оказались средние значения ИПЦ за три месяца (t_mean_lag3), значение ИПЦ с лагом 1 ($t-1$), номер месяца ($month$) и стандартное отклонение валютного курса за последние 6 месяцев ($curs_std_lag6$).

Рис. 9. Значимость факторов модели для Новосибирской области методом эластичной сети



Источник: расчеты авторов.

Таким образом, можно сделать вывод о возможности использования методов машинного обучения для прогнозирования инфляции в макрорегионе Сибирь и регионах Сибири совместно с традиционными эконометрическими методами. Вместе с тем для каждого региона необходима предварительная оценка качества прогноза методами машинного обучения и целесообразности их использования в сравнении с традиционными эконометрическими методами.

6. ЗАКЛЮЧЕНИЕ

По результатам проведенного исследования можно сделать вывод, что качество прогнозов инфляции в макрорегионе Сибирь и регионах Сибири с помощью методов машинного обучения сопоставимо с традиционными эконометрическими методами ARIMA и ARIMAX, но для каждого региона необходима предварительная оценка качества прогноза и целесообразности их использования. Для большинства сибирских регионов и для макрорегиона Сибирь на сроках свыше года методы машинного обучения показывают качество лучше бенчмарков, в отличие от прогнозов на срок до года. Прогнозирование методом комбинирования моделей машинного обучения в большинстве случаев предпочтительнее, чем каким-то одним.

Возможным направлением развития исследования может стать переход от прогнозирования в режиме псевдореального времени к реальному. В этом случае мы сможем основываться на сценарных прогнозах номинального курса рубля и добавлять сценарные прогнозы других факторов. Оценка качества прогноза компонентов инфляции

и сопоставление качества прогнозов до и после кризисных явлений в экономике также представляет отдельный интерес для исследования.

Кроме того, к полученному прогнозу можно добавлять экспертную корректировку, которая поможет учесть не проявившие себя в прошлом разовые факторы, но способные повлиять на уровень региональной инфляции.

По мере необходимости возможно расширение изложенного подхода на другие российские регионы при сохранении индивидуального учета региональной специфики и экспертной корректировки полученных результатов.

7. СПИСОК ЛИТЕРАТУРЫ

1. Байбуза И. Прогнозирование инфляции с помощью методов машинного обучения // Деньги и кредит. – 2018. – № 4. – С. 42–59.
2. Павлов Е. Прогнозирование инфляции в России с помощью нейронных сетей // Деньги и кредит. – 2020. – № 1. – С. 57–73.
3. Яковлева К. Оценка экономической активности на основе текстового анализа // Деньги и кредит. – 2018. – № 4. – С. 26–41.
4. Athey S., Imbens G.W. Machine Learning Methods Economists Should Know About // Annual Review of Economics. – 2019. – № 11. – P. 685–725.
5. Barkan O., Benchimol J., Caspi I., Hammer A., Koenigstein N. Forecasting CPI Inflation Components with Hierarchical Recurrent Neural Networks // Bank of Israel Working Paper. – 2021.
6. Chakraborty C., Joseph A. Machine Learning at Central Banks // Bank of England Working Paper. – 2017. – № 674.
7. Diebold F.X., Mariano R.S. Comparing Predictive Accuracy // Journal of Business and Economic Statistics. – 1995. – № 13 (3). – P. 253–263.
8. Jung J., Patnam M., Ter-Martirosyan A. An Algorithmic Crystal Ball: Forecast-based on Machine Learning // IMF Working Paper. – 2018. – № 18/230.
9. Mamedli M., Shubitov D. Forecasting Russian CPI with Data Vintages and Machine Learning Techniques // The Bank of Russia Working Paper Series. – 2021. – № 70.
10. Ozgur O., Akkoç U. Inflation Forecasting in an Emerging Economy: Selecting Variables with Machine Learning Algorithms // International Journal of Emerging Markets. – 2021.