



**ЕФСР**

ЕВРАЗИЙСКИЙ ФОНД  
СТАБИЛИЗАЦИИ И РАЗВИТИЯ

# **Алгоритмы машинного обучения для краткосрочного прогнозирования темпов прироста реального ВВП**

Москва,  
декабрь 2024 г.

EFSD.ORG

# Цель и задачи исследования

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

*Если мы стремимся максимально эффективно использовать данные для поиска ответов на вопросы, то необходимо отказаться от зависимости от традиционных моделей и расширить набор инструментов.*

## Задачи исследования:

- 1 изучить возможности методов и алгоритмов машинного обучения для краткосрочного прогнозирования макроэкономических показателей,
- 2 разработать модели машинного обучения с использованием Python,
- 3 сравнить прогностические свойства традиционных моделей (AR, VAR, BVAR) и методов/алгоритмов машинного обучения.

# 1. ОБЗОР ИСПОЛЬЗУЕМЫХ МЕТОДОВ

# Используемые методы прогнозирования

## Традиционные эконометрические модели

- Модель одномерной авторегрессии (AR)
- Модель векторной авторегрессии (VAR)
- Байесовская модель векторной авторегрессии (BVAR)

## Классические алгоритмы с учителем

- Ridge-регрессия
- Lasso-регрессия
- Elastic Net

## Ансамблевые методы

- Bagging
- Random Forest
- Boosting

## Нейронные сети

- Многослойный перцептрон

Методы и алгоритмы  
машинного обучения

# Традиционные эконометрические модели

## Модель одномерной авторегрессии (AR)

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

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

где  $\phi_0, \phi_1$  – параметры,

$\varepsilon_t$  – случайные ошибки модели.

## Модель векторной авторегрессии (VAR)

Традиционная VAR-модель без ограничений имеет вид:

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

где  $Y_t$  – вектор исходных переменных,

$c$  – вектор постоянных коэффициентов,

$A_1, \dots, A_p$  – матрицы оцененных параметров для различных временных лагов  $l = 1, 2, \dots, p$ ,

$\varepsilon_t$  – вектор случайных ошибок модели.

## Байесовская модель векторной авторегрессии (BVAR)

BVAR является моделью с ограничениями (в сравнении с традиционной VAR-моделью). В рамках BVAR-модели при увеличении временного лага объясняющих показателей оцениваемые параметры обнуляются,

# Классические алгоритмы с учителем

$$\Delta \ln rGDP_t = \beta_0 + \beta_1 \Delta \ln ip_{t-1} + \beta_2 \Delta \ln inv_{t-1} + \dots + \beta_{17} \Delta fdepir_{t-1} + \varepsilon_t$$

## Ridge-регрессия

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

Контролирует величину сжатия

## Lasso-регрессия

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

## Elastic Net

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

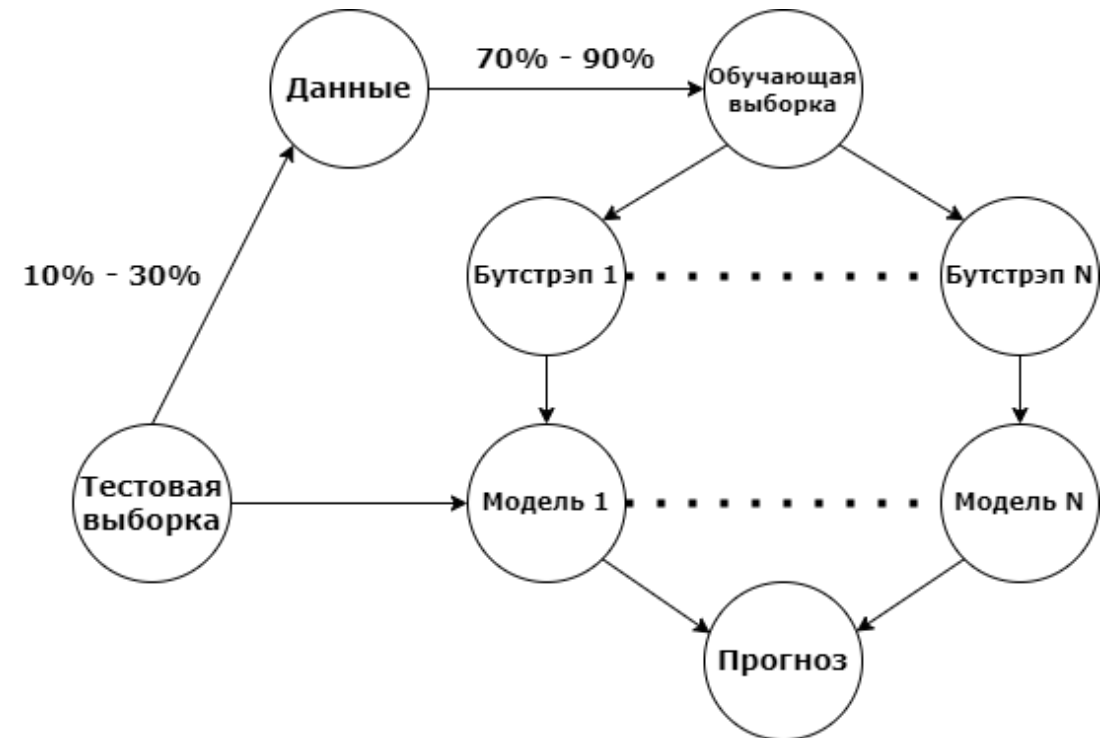
Вес пенализации

# Ансамблевые методы: Bagging

Пусть набор обучающих данных  $D = \{\mathbf{x}_n, y_n\}_{n=1}^N$ , а тестовая выборка  $B = \{\mathbf{x}_n\}_{t=1}^T$ .

**Алгоритм Bagging включает следующие шаги:**

1. для  $b = 1, 2, \dots, B$  делается бутстрэп выборка с заменой  $D_b^*$ ;
2. для данной бутстрэп выборки строится дерево решений  $B_b^*$ ;
3. для периода  $t = 1, 2, \dots, T$  и для каждой из бутстрэп выборок  $b = 1, 2, \dots, B$  рассчитывается  $\hat{y}_{tb}^*$ ;
4. на основе  $\hat{y}_{tb}^*$  рассчитывается среднее значение  $\hat{y}_t$ .



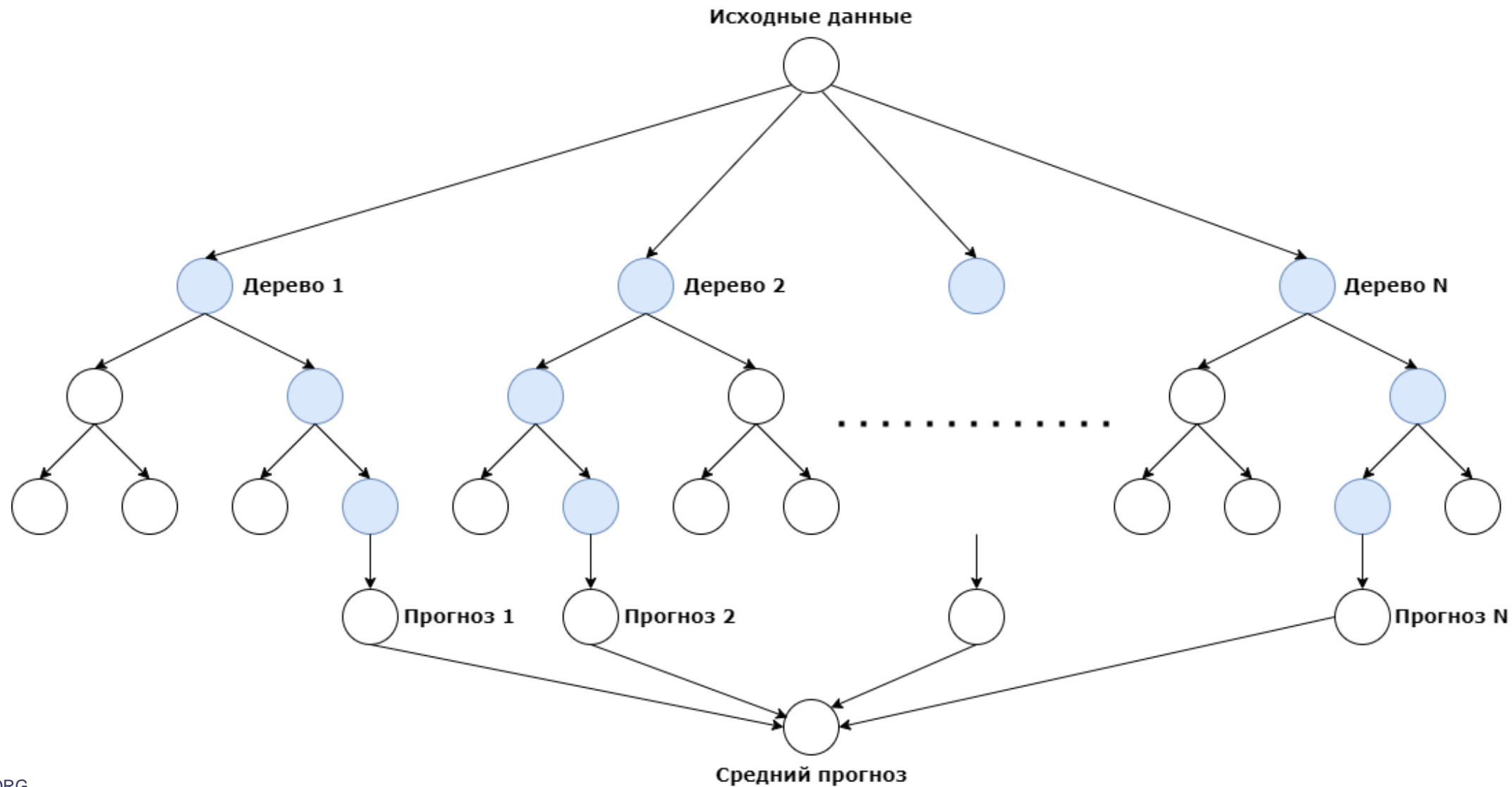
# Ансамблевые методы: Random Forest

Пусть набор обучающих данных  $D = \{\mathbf{x}_n, y_n\}_{n=1}^N$ , а тестовая выборка  $\mathcal{B} = \{\mathbf{x}_n\}_{t=1}^T$ .

**Алгоритм Random Forest включает следующие шаги:**

1. для  $b = 1, 2, \dots, B$  делается бутстреп выборку с заменой  $D_b^*$ ;
2. для данной бутстреп выборки строится дерево решений  $\mathcal{B}_b^*$ . В отличие от Bagging дерево решений строится для  $C$  предикторов, где  $C < D$ ;
3. для периода  $t = 1, 2, \dots, T$  и для каждой из бутстрэп выборок  $b = 1, 2, \dots, B$  рассчитывается  $\hat{y}_{tb}^*$ ;
4. на основе  $\hat{y}_{tb}^*$  рассчитывается среднее значение  $\hat{y}_t$ .

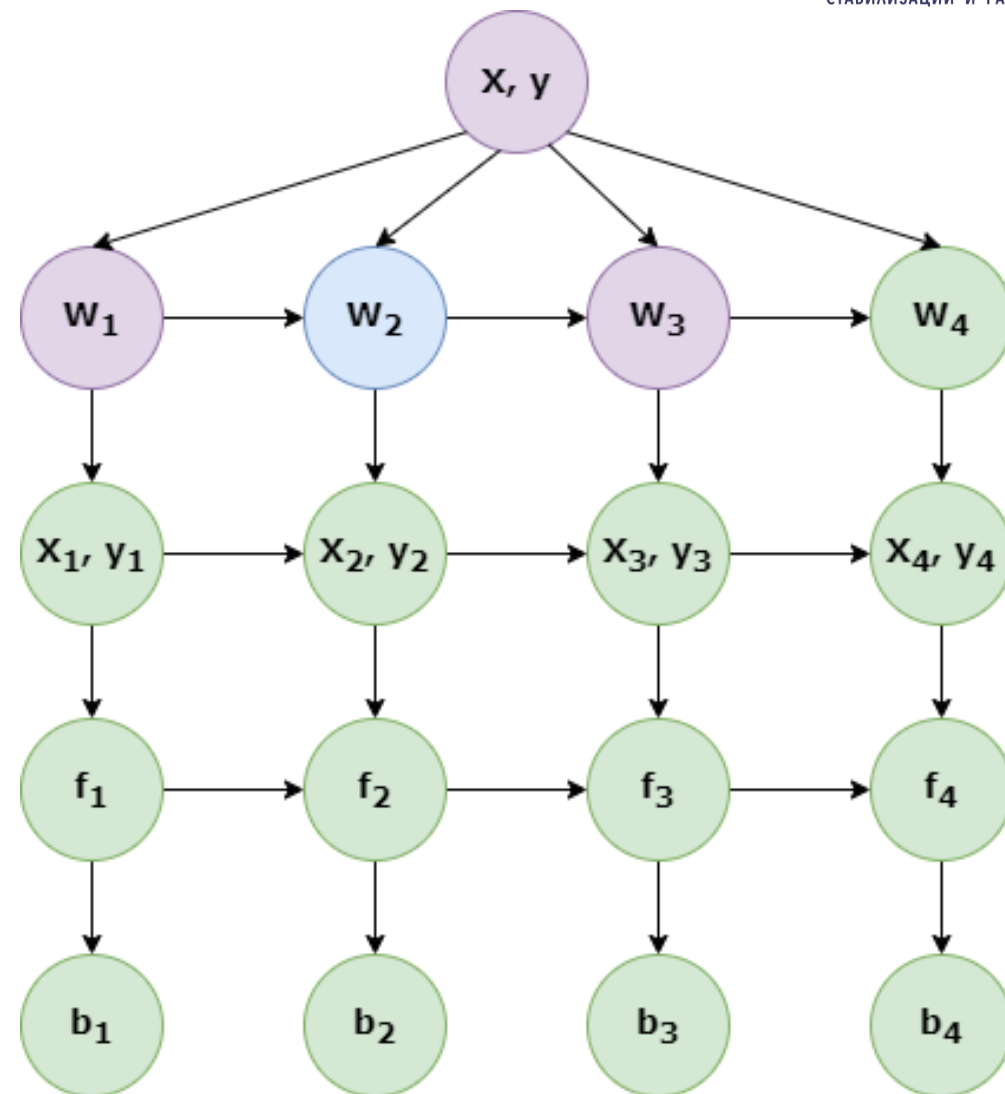
# Ансамблевые методы: Random Forest



# Ансамблевые методы: Boosting

Алгоритм Boosting включает следующие шаги:

1. Задаются исходные значения весов  $w_n^1$ , где  $n = 1, 2, \dots, N$  ( $N$  – количество строк обучающей выборки).
2. Задается общее количество бутстрэп повторений (например,  $T = 100$ ).
3. Для каждого бутстрэп повторения строится дерево решений и прогнозируется среднее значение зависимой переменной на концах дерева (листьях) на основе обучающей выборки.



# Ансамблевые методы: Boosting

4. Рассчитывается вектор ошибок наблюдения  $L_n^t$  для  $n = 1, 2, \dots, N$  согласно формулам:

$$D^t = \underbrace{\max}_n \{|y_n - \hat{y}_n|\},$$

$$L_n^t = \frac{|y_n - \hat{y}_n|}{D^t}.$$

5. Рассчитываем среднее значение ошибок наблюдения:

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

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

# Ансамблевые методы: Boosting

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

7. Обновляются весовые коэффициенты модели по следующему правилу

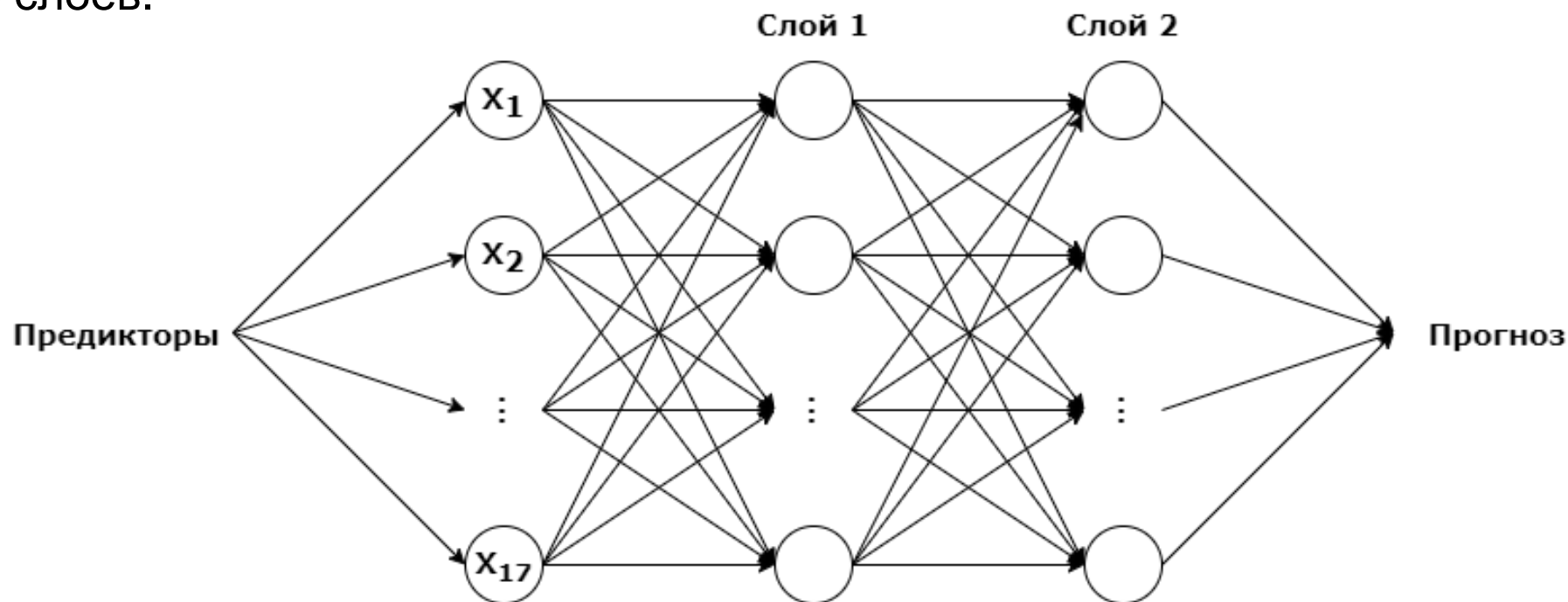
$w_n^{t+1} = \frac{w_n^t (\beta^t)^{1-L_n^t}}{\sum_{n=1}^N w_n^t (\beta^t)^{1-L_n^t}}$ , что уменьшает вес наблюдений со сравнительно большей ошибкой.

8. В случае  $\bar{L}^t \geq 0,5$  рассчитываются значения прогнозов зависимой переменной как взвешенную медианную величину с помощью весов  $\log\left(\frac{1}{\beta^t}\right)$ .

# Нейронные сети

Используется **нейронная сеть с прямой связью (feed-forward neural networks)**.

Например, имеем обучающую выборку  $\{x_n, y_n\}_{n=1}^N$ . Сеть начинается с входного слоя, состоящего из вектора предикторов  $(x_0, x_1, x_2, \dots, x_{17})$ . Далее сеть проходит через один или несколько слоев.



# Нейронные сети

Первый скрытый слой является функцией входного слоя, а каждый последующий слой является функцией последнего.

Согласно диаграмме сеть переходит из последнего скрытого слоя в выходной слой, представляющий собой целевую переменную или переменные.

Каждый скрытый слой состоит из нейронов, показанных в кружках. На диаграмме входной нейрон имеет 17 нейронов.

Нейроны первого скрытого слоя обозначены  $z_0^{(1)}, z_1^{(1)}, z_2^{(1)}$  и т.д., нейроны второго скрытого слоя -  $z_0^{(2)}, z_1^{(2)}, z_2^{(2)}$  и т.д.

Обратим внимание на линии, которые идут от нейронов входного слоя к нейронам первого скрытого слоя.

Как только нейроны первого скрытого слоя заданы они становятся предикторами для следующего по очередности слоя. Наконец, когда нейроны в последнем скрытом слое фиксированы они действуют как предикторы для выходного слоя.

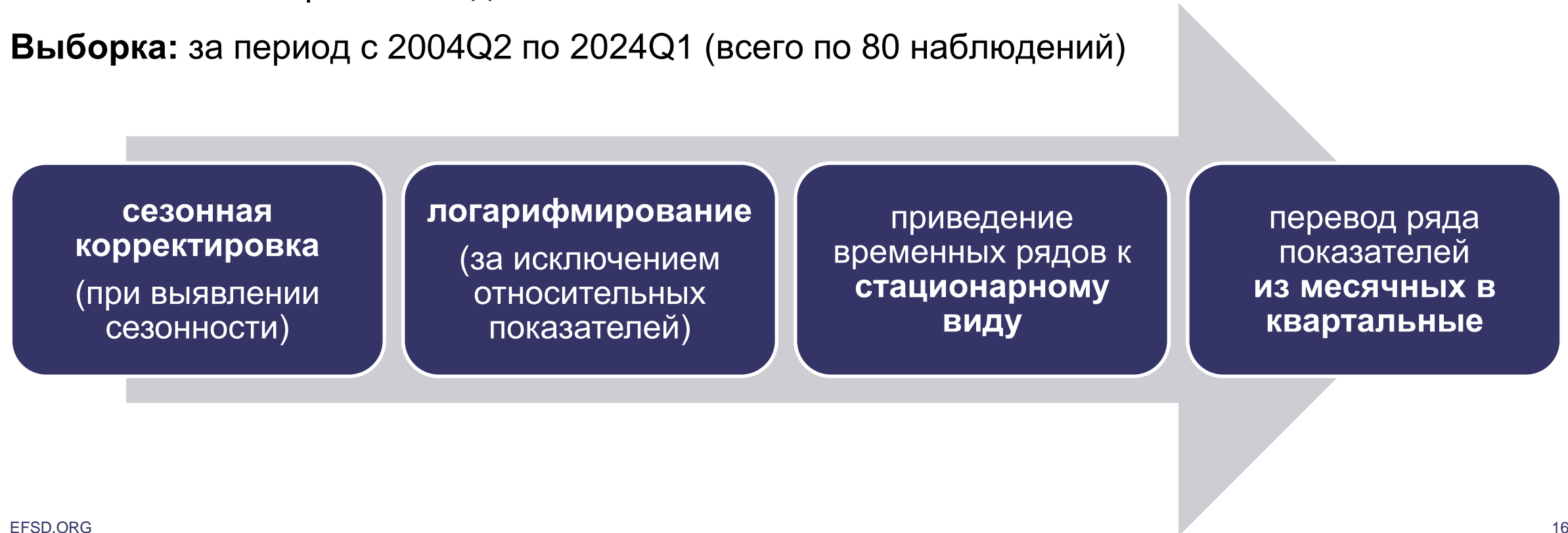
## **2. ИСПОЛЬЗУЕМЫЕ ДАННЫЕ**

# Используемые данные

**Тип данных:** макроэкономические показатели – 1 прогнозируемый показатель и 17 потенциальных предикторов

**Частотность:** квартальные данные

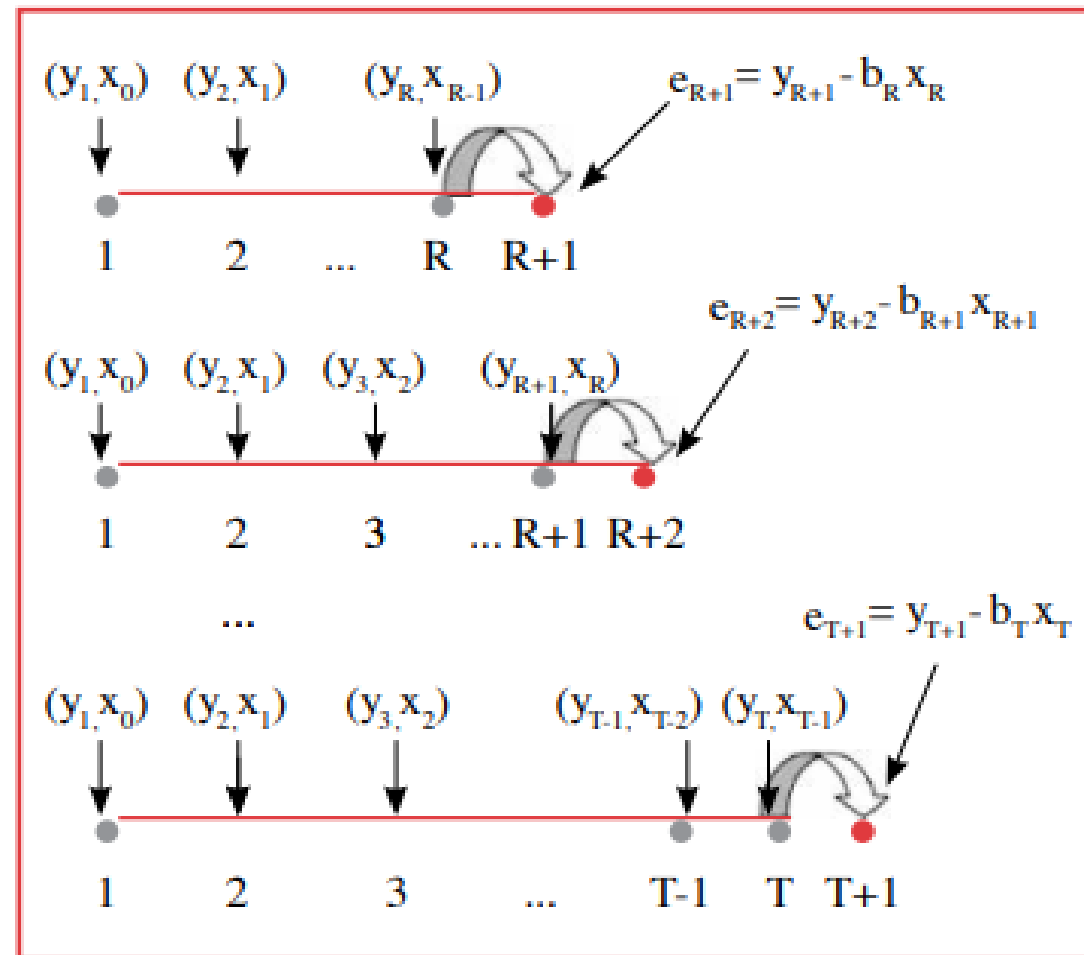
**Выборка:** за период с 2004Q2 по 2024Q1 (всего по 80 наблюдений)



## **3. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ**

# Рекурсивная схема регрессионного эксперимента

1. Вся выборка первоначально делится на 2 части – обучающую (первые 70% наблюдений) и тестируемую (последние 30% наблюдений).
2. **Оценивание модели** на обучающей выборке.
3. Используя оцененную модель, **производится прогноз** на 1 квартал вперед.
4. **Увеличивается длина обучающей выборки** на 1 квартал и заново оцениваются параметры модели.
5. Используя **переоцененную модель**, **производится прогноз** на 1 квартал вперед.

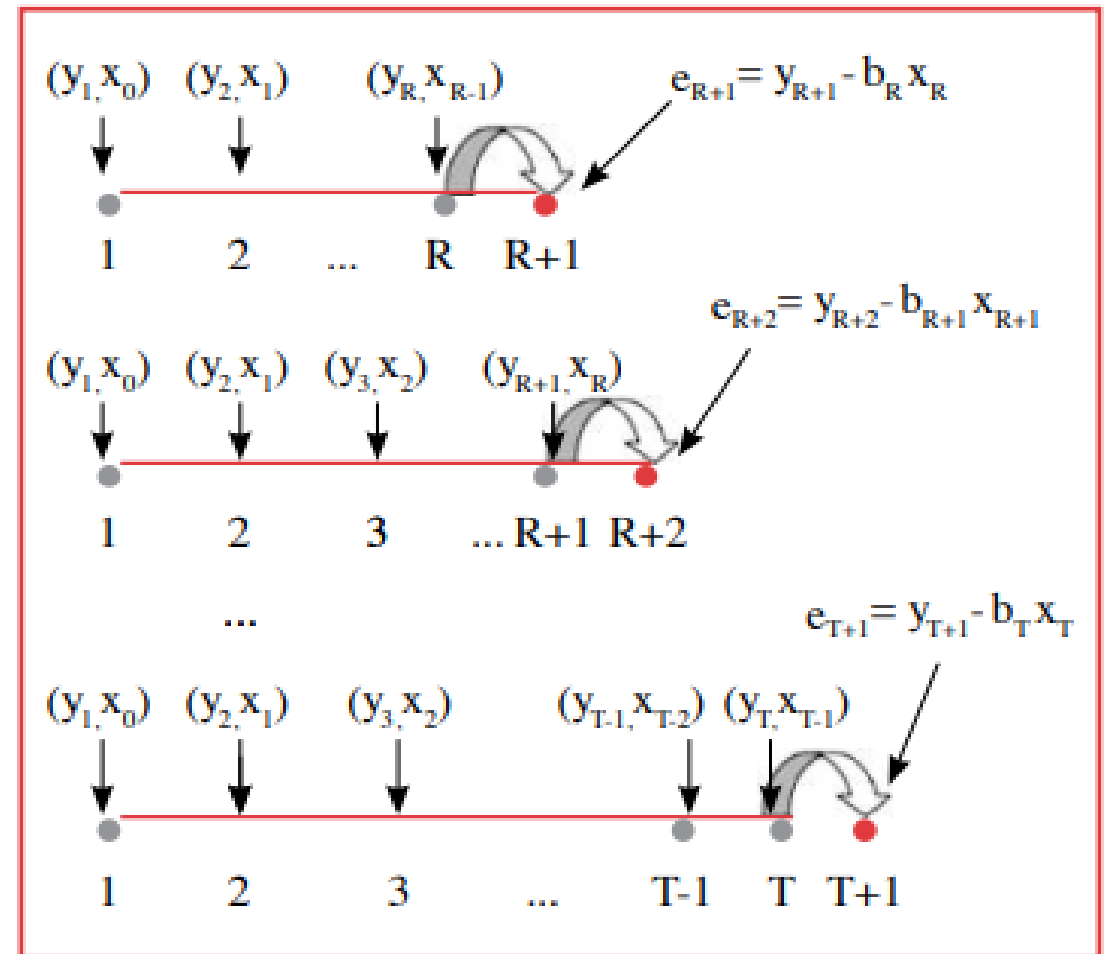


# Рекурсивная схема регрессионного эксперимента

6. Аналогично шагам 4 и 5, происходит увеличение выборки, переоценка модели и построение прогноза на один квартал вперед до тех пор, пока выборка не достигнет периода с 2018Q2 по 2023Q4.
7. В результате полученные **24 точечных прогноза** темпов прироста ВВП сравниваются с фактическими данными и **рассчитываются два индекса**, которые являются функциями потерь (*loss functions*):

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

$$MAD_i = \frac{1}{n} \sum_{t=1}^T |y_t - \hat{y}_t|$$



# Результаты экспериментов

AR-модель по качеству прогноза превосходит VAR- и BVAR-модели.

Большинство рассматриваемых методов и алгоритмов машинного обучения по качеству прогноза превосходят AR-модель.

Согласно статистике Диболда-Мариано, различия между прогнозами AR-модели и иных анализируемых методов (за исключением VAR-модели) не существенны.

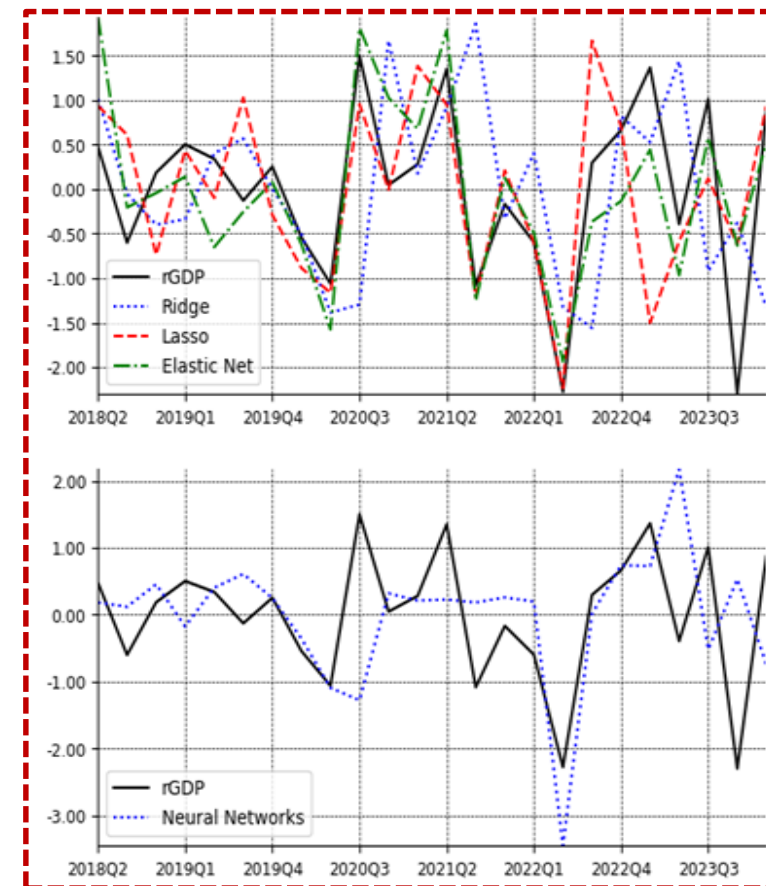
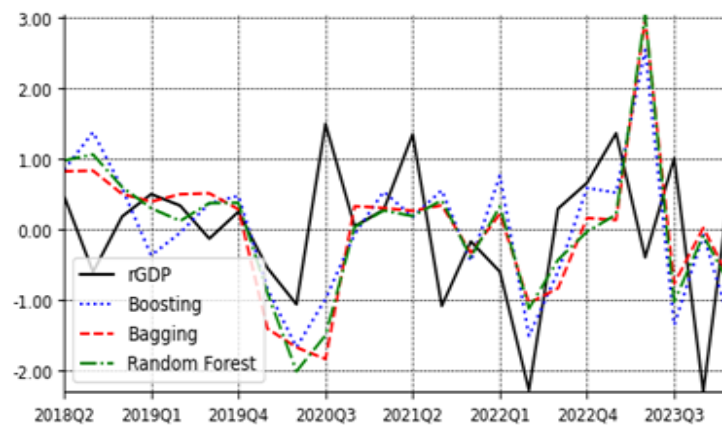
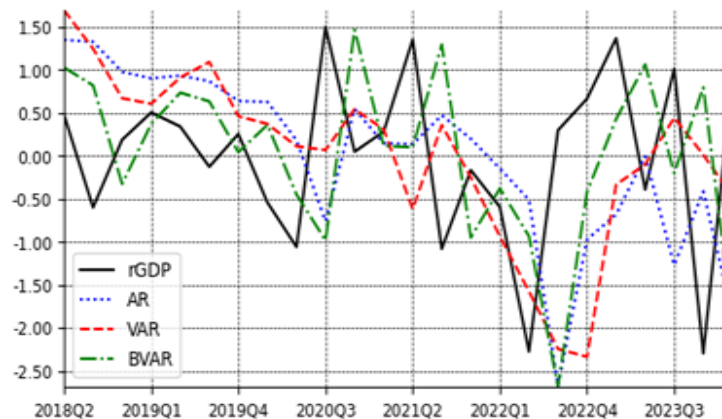
| Модели         | RMSE  | RMSE (по отнош. к AR) | DM-stat. | MAD   | MAD (по отнош. к AR) | DM-stat. |
|----------------|-------|-----------------------|----------|-------|----------------------|----------|
| AR             | 2,121 | --                    | --       | 1,601 | --                   |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

\* существенна на 5% уровне значимости.

# Результаты экспериментов

При использовании моделей машинного обучение **не наблюдается лагового сдвига.**

**Пики и спады** темпов прироста реального ВВП наиболее точно прогнозируются с помощью **методов регуляризации (Ridge, Lasso, Elastic Net) и Neural Networks.**



# Результаты экспериментов

Все методы взвешивания обеспечивают более низкие значения  $RMSE$  и  $MAD$ . При этом **наиболее высокая точность** наблюдается при взвешивании **с помощью МНК весов**

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

\* существенна на 5% уровне значимости.

# Результаты экспериментов

Наиболее точное улавливание спадов и пиков темпов прироста реального ВВП – в случае комбинированного прогноза **по МНК весам**.

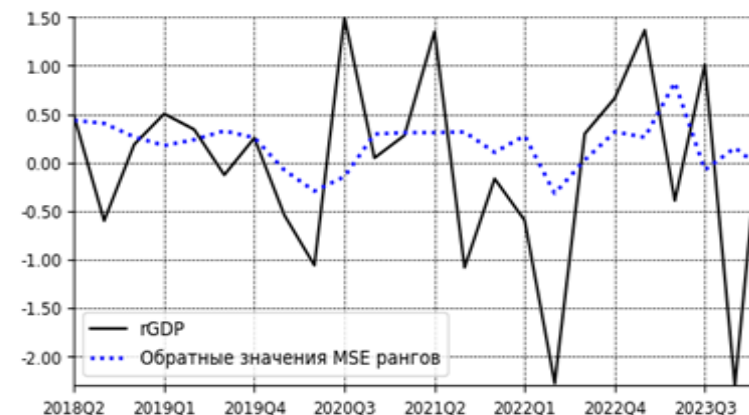
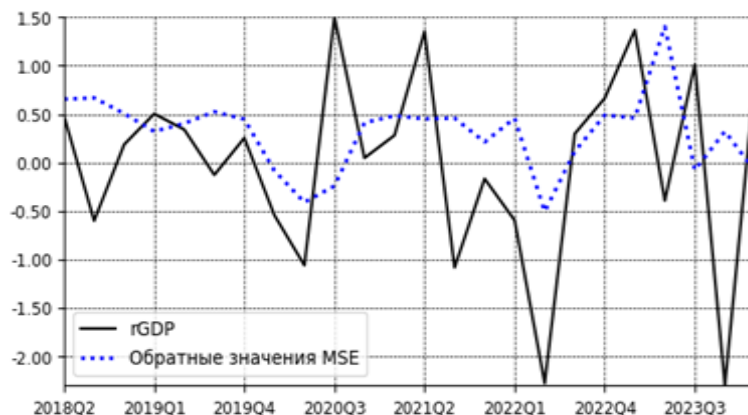
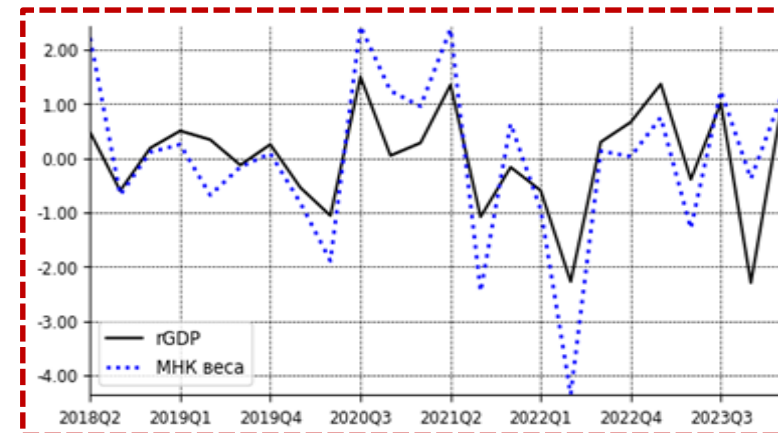
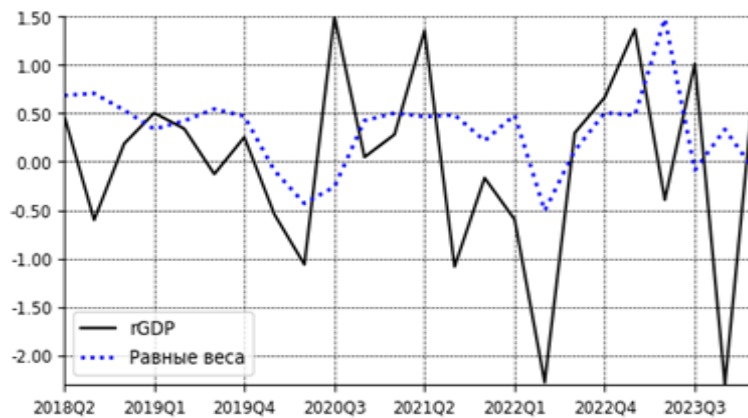


Рисунок. Комбинированные прогнозы темпов прироста реального ВВП

# Заключение

1. **Большинство** рассматриваемых в работе **методов и алгоритмов машинного обучения** по точности прогноза **превосходит традиционные инструменты.**
2. Согласно результатам статистики Диболда-Мариано, **различия между прогнозами** анализируемых методов **статистически не существенны.**
3. При использовании методов и алгоритмов машинного обучения **не наблюдается лагового сдвига, характерного для AR-, VAR- и BVAR-моделей.**
4. **Результаты прогнозов существенно улучшаются** за счет применения **комбинированного прогноза.** Наибольшая точность – на основе МНК весов.



Շնորհակալութիւնն ուշադրութեան համար.

Дзякуй за ўвагу!

Назар аударғандарыңызға рахмет!

Назар салганыңыздарга рахмат!

Благодарю за внимание!

Раҳмати калон!

**Тарас Цукарев**

Начальник отдела Группы  
главного экономиста ПБ ЕФСР  
[ttsukarev@efsd.org](mailto:ttsukarev@efsd.org)

**Карен Погосян**

Ст. экономист Группы главного  
экономиста ПБ ЕФСР  
[kpoghosyan@efsd.org](mailto:kpoghosyan@efsd.org)

**Кирилл Лемба**

Эксперт Группы главного  
экономиста ПБ ЕФСР  
[KLemba@efsd.org](mailto:KLemba@efsd.org)