

Научная статья

УДК 338

DOI: <https://doi.org/10.18721/JE.17604>



## ВЫЯВЛЕНИЕ ФАКТОРОВ ВОЗДЕЙСТВИЯ НА СЕКТОР СВЯЗИ И ТЕЛЕКОММУНИКАЦИЙ С ПРИМЕНЕНИЕМ АНСАМБЛЕВЫХ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

И.Р. Бадыкова , К.Р. Биктимирова

Казанский национальный исследовательский технологический университет,  
г. Казань, Российская Федерация

 [idelia.badykova@gmail.com](mailto:idelia.badykova@gmail.com)

**Аннотация.** *Актуальность.* Сектор связи и телекоммуникаций играет одну из ключевых ролей в развитии экономики. В связи с этим особенно важным становится определение факторов, оказывающих на него воздействие. При исследовании российского рынка является существенным учет неоднородности развития регионов. *Целью* исследования выступает моделирование деятельности сектора телекоммуникаций и связи и выявление факторов, которые оказывают воздействие на него, с учетом региональной специфики. *Методы.* Эмпирическую базу данного исследования составили данные официальной статистической отчетности по 85 субъектам Российской Федерации поквартально за 2017–2023 гг. В качестве результирующей переменной взяты доходы от услуг связи, независимых переменных – объем информации, переданной от/к абонентам при доступе в интернет; индексы тарифов на услуги связи для юридических лиц; среднесписочная численность работников в сфере телекоммуникаций; базовый индекс потребительских цен на товары и услуги; среднемесячная заработная плата работников в экономике; почтовые переводы. С использованием языка программирования Python построены модели линейной регрессии, Ridge, Bagging, RandomForestRegressor, GradientBoostingRegressor и XGBRegressor. Для интерпретации результатов ансамблевых методов использован метод SHAP. *Результаты.* Анализ подтверждает неравномерное развитие сектора связи и телекоммуникаций в рамках субъектов страны. Среди построенных моделей наилучший результат достигнут с применением GradientBoostingRegressor, а анализ SHAP-значений выявил влияние на доходы от услуг связи различных факторов, таких как повышение среднемесячной зарплаты, увеличение объема информации (трафика), индексы тарифов услуг связи для юридических лиц, а также негативное влияние на них базового индекса потребительских цен и незначительное – почтовых переводов. *Выводы.* Выявление факторов, оказывающих воздействие на сектор связи и телекоммуникаций в условиях российской экономики, является важным и имеет как теоретическую, так и практическую значимость. Так, полученные результаты могут быть использованы для различных целей: стратегического планирования, оптимизации инвестиций, государственного регулирования, маркетинга и продаж, прогнозирования и планирования. *Направления дальнейших исследований.* Несмотря на то, что качество модели с учетом количества рассмотренных наблюдений представляется вполне высоким, в полученных моделях можно видеть некоторое переобучение, для устранения которого с целью проведения дальнейших исследований целесообразно при генерировании выборки добавлять значения наблюдений за новые периоды. Кроме того, целесообразным представляется генерировать нейронные сети для получения еще более точных оценок влияния различных факторов на сектор связи и телекоммуникаций, а также для получения более точных прогнозных оценок.

**Ключевые слова:** сектор связи и телекоммуникаций, факторы воздействия, машинное обучение, ансамблевые методы, регионы

**Для цитирования:** Бадыкова И.Р., Биктимирова К.Р. (2024) Выявление факторов воздействия на сектор связи и телекоммуникаций с применением ансамблевых методов машинного обучения. П-Economy, 17 (6), 61–78. DOI: <https://doi.org/10.18721/JE.17604>



## IDENTIFYING IMPACT FACTORS ON THE COMMUNICATIONS AND TELECOMMUNICATIONS SECTOR USING ENSEMBLE MACHINE LEARNING METHODS

I.R. Badykova  , K.R. Biktimirova

Kazan National Research Technological University, Kazan, Russian Federation

 [idelia.badykova@gmail.com](mailto:idelia.badykova@gmail.com)

**Abstract.** *Relevance.* The communications and telecommunications sector plays one of the key roles in economic development. In this regard, it is especially important to determine the factors that influence this sector. When studying the Russian market, it is essential to take into account the heterogeneity of regional development. *The purpose of the study* is to model the activities of the communications and telecommunications sector and identify the factors that influence it, taking into account regional specifics. *Methods.* The empirical basis of this study is formed by the data of official statistical reporting for 85 constituent entities of the Russian Federation quarterly for 2017–2023. The resulting variable is revenue from communication services, the independent variables are the volume of information transmitted from/to subscribers when accessing the Internet; tariff indices for communication services for legal entities; average number of employees in the telecommunications sector; basic consumer price index for goods and services; average monthly wages of employees in the economy; postal transfers. Using the Python programming language, the following models were built: linear regression, Ridge, Bagging, RandomForestRegressor, GradientBoostingRegressor and XGBRegressor. The SHAP method was used to interpret the results of ensemble methods. *Results.* The analysis confirms the uneven development of the communications and telecommunications sector within the country's regions. Among the constructed models, the best result was achieved using GradientBoostingRegressor, and the analysis of SHAP indicators revealed the influence of various factors on revenues from communications services, such as an increase in the average monthly salary, an increase in the volume of information (traffic), tariff indices for communications services for legal entities, as well as the negative impact of the basic consumer price index and an insignificant impact of postal transfers. *Conclusions.* Identifying factors that affect the communications and telecommunications sector in the context of the Russian economy is important and has both theoretical and practical significance. Thus, the results obtained can be used for various purposes: strategic planning, investment optimization, government regulation, marketing and sales, forecasting and planning. *Directions for further research.* Despite the fact that the quality of the model, taking into account the number of observations considered, seems quite high, some overfitting can be seen in the obtained models, to eliminate which, for the purpose of conducting further research, it is advisable to add observation values for new periods when generating a sample. In addition, it seems advisable to generate neural networks to obtain even more accurate estimates of the impact of various factors on the communications and telecommunications sector, as well as to obtain more accurate forecast estimates.

**Keywords:** communications and telecommunications sector, impact factors, machine learning, ensemble methods, regions

**Citation:** Badykova I.R., Biktimirova K.R. (2024) Identifying impact factors on the communications and telecommunications sector using ensemble machine learning methods. *π-Economy*, 17 (6), 61–78. DOI: <https://doi.org/10.18721/JE.17604>



## Введение

В условиях современности сектор связи и телекоммуникаций становится стратегически важным. О том, что его роль является одной из ключевых в экономике, говорят многие исследователи, включая таких, как Х.Т. Омбре с соавторами [1], Х.Л. Гомес-Барросо и Р. Марбан-Флорес [2], З. Латиф с соавторами [3], М. Кочча [4], Л. Хитт [5], Р. Гхолами с соавторами [6], Е.А. Фокина [7] и др.

Стремительный рост технологий, включая развитие искусственного интеллекта, растущий спрос на услуги связи, текущая политическая и экономическая обстановка и многие другие факторы оказывают воздействие на сектор, и становится важным понять, что из себя представляют детерминанты его развития именно для российского рынка. Учитывая неоднородное развитие регионов, представляется интересным провести исследование в разрезе субъектов Российской Федерации, которые и выступают объектом исследования. Предметом является непосредственно сектор связи и телекоммуникаций.

Целью данного исследования является проведение анализа развития сектора связи и телекоммуникаций и выявление факторов, которые оказывают воздействие на него, с учетом региональной специфики.

Для достижения поставленной цели должны быть выполнены следующие задачи:

- изучение теоретических аспектов развития сектора связи и телекоммуникаций;
- сбор и систематизация статистических данных;
- использование аналитического инструментария для эконометрического моделирования данных;
- выявление лучшей модели и интерпретация полученных результатов.

Научная новизна исследования состоит в построении модели для выявления факторов, оказывающих воздействие на развитие сектора связи и телекоммуникаций, с отведением важной роли региональной специфике. Соответственно, при моделировании учтено неравномерное развитие экономики в различных регионах страны. Кроме того, использованы актуальные методы эмпирического исследования, а именно методы машинного обучения, что позволяет получать достаточно точные модели и делать обоснованные выводы о характере влияния факторов на развитие сектора связи и телекоммуникаций.

## *Литературный обзор*

Вопрос о том, какие факторы оказывают влияние на развитие сектора телекоммуникаций и связи в экономике, является достаточно широко обсуждаемым как в зарубежной, так и отечественной научной литературе.

Изменения в экономике, различного рода рыночные колебания, как правило, существенно влияют на развитие сектора. Безусловно, одним из наиболее важных факторов его развития также является внедрение инноваций [8–10]. Непрерывное развитие информационно-коммуникационных технологий оказывает существенное положительное воздействие.

Важной особенностью сектора выступает монополизация [11, 12], приводящая к тому, что государство занимает регулируемую позицию. Кроме того, очевидно, что все стратегии и действия той или иной компании не должны выходить за рамки имеющегося законодательства и его изменений [13].

Интересным является исследование несбалансированной панели Р. Кацем и Х. Юнгом. Ученые на основании анализа данных 108 стран за период 2009–2018 гг. выявили негативное влияние увеличения регуляторных сборов, налога на прибыль и акцизов как на инвестиции в телекоммуникационный сектор, так и на цены на услуги [14].

В. Йиланци и Э.Н. Килчи, проведя эмпирическое исследование для турецкого рынка, пришли к выводу о наличии долгосрочной взаимосвязи между экономическим ростом, иностранными инвестициями, открытостью торговли и телекоммуникационной инфраструктурой [15].

Л.В. Улезлова проводит классификацию факторов. Так, ученый говорит о наличии положительных и отрицательных факторов. К первой группе она относит развитие сотрудничества в

сфере телекоммуникаций и связи, потребность в создании новых технологий; ко второй – неравномерное развитие регионов страны, климатические условия, неравномерное распределение населения по территории, ограниченную покупательную способность [16].

М.Ю. Малкина, в свою очередь, эмпирически обнаружила существенное негативное влияние нефтяных шоков в период пандемии COVID19 [17].

Безусловно, крайне важным фактором для российского рынка телекоммуникаций и связи оказалось введение санкций. Так, в условиях зависимости российской информационной инфраструктуры от наличия поставок зарубежного оборудования российским предприятиям приходится быстро реагировать и подстраиваться под текущие обстоятельства. Среди решений проблемы, которые применяют компании, – смена поставщиков, параллельный импорт, импортозамещение. На сегодняшний день наблюдается достаточно продолжительный временной лаг между разработкой и внедрением методов информационной безопасности [18, 19]. В случае с действующими в условиях высокой неопределенности операторами связи возникают вопросы относительно развертывания сетей 5G [20–23].

Не умаляя значимости работ вышеупомянутых авторов, следует заметить, что в контексте исследования российского рынка крайне важным является понимание неоднородности развития субъектов федерации, исходя из чего мы и построим эмпирическую часть исследования.

### Методы и материалы

Эмпирическую базу данного исследования составили данные официальной статистической отчетности по 85 субъектам Российской Федерации поквартально за 2017–2023 гг., собранные Федеральной службой государственной статистики (далее – Росстат). Таким образом, количество наблюдений в рамках панели равно 2380.

В качестве результирующей переменной взяты доходы от услуг связи, поскольку этот показатель представляет собой ключевой аспект для анализа и отражает экономическое положение отрасли, включая выручку от мобильной связи, широкополосного интернета, фиксированной связи и других видов услуг. Именно данный показатель предоставляет возможности для оценки динамики спроса на услуги связи, конкурентоспособности операторов связи, эффективности инвестиций в развитие инфраструктуры и технологий, а также влияния регулирующих мер и законодательства на финансовые показатели отрасли.

Отбор независимых переменных объясняется наличием данных для всех регионов в рамках базы данных Росстат. Таким образом, в рамках модели рассматриваются следующие факторы: объем информации, переданной от/к абонентам при доступе в интернет; индексы тарифов на услуги связи для юридических лиц; среднесписочная численность работников в сфере телекоммуникаций; базовый индекс потребительских цен (БИПЦ) на товары и услуги; среднемесячная заработная плата работников в экономике; почтовые переводы.

Соответственно, гипотезами нашего исследования выступили:

*H<sub>1</sub>: Объем информации, переданной от/к абонентам при доступе в интернет, оказывает положительное влияние на результирующий показатель.*

*H<sub>2</sub>: Численность работников телекоммуникаций положительно сказывается на показателе доходов от услуг связи.* Увеличение персонала позволяет улучшить качество обслуживания и, как следствие, привести к росту доходов.

*H<sub>3</sub>: Индексы тарифов услуг связи для юридических лиц оказывают положительное воздействие на зависимую переменную исследования.* Можно предположить, что клиенты склонны тратить больше при увеличении цен.

*H<sub>4</sub>: Почтовые переводы положительно влияют на доходы от услуг связи.* Почтовые переводы предоставляют дополнительный источник дохода для почтовых служб, увеличивая общий объем операций и привлекая новых клиентов.



$H_5$ : Средняя заработная плата по региону положительно влияет на результирующую переменную. Чем выше зарплата, тем большие средства клиенты могут тратить на услуги связи.

$H_6$ : Базовый индекс потребительских цен оказывает негативное влияние на доходы от услуг связи. Можно предположить, что при росте цен на потребительские товары и услуги клиенты могут проявлять большую осторожность в расходовании денег на дополнительные услуги связи.

Учитывая достаточно большой объем выборки, целесообразным представляется использование для проведения статистического анализа, визуализации данных и построения регрессионных и ансамблевых моделей языка программирования Python, позволяющего глубоко исследовать взаимосвязи в данной отрасли. Python является гибким, мощным инструментом, обладающим большим количеством библиотек для анализа данных и прогнозирования.

Можно выделить следующие этапы нашего исследования:

1. Описательная статистика данных. Прежде всего, в рамках исследования использована библиотека Pandas, позволяющая обрабатывать и анализировать табличные данные в формате DataFrame. Кроме того, для визуализации данных построены гистограммы и бокс-плоты («ящики с усами») для каждой из рассматриваемых переменных.

2. Построение матрицы корреляций.

3. Применение классических методов статистического анализа и моделей машинного обучения. Прогнозирование доходов от услуг связи имеет важное экономическое и стратегическое значение для сферы телекоммуникаций. Для исследования моделей прогнозирования применяются линейная регрессия, Ridge, Bagging, RandomForestRegressor, GradientBoostingRegressor и XGBRegressor и осуществляется сравнительный анализ качества полученных моделей. Каждая из них имеет свои уникальные особенности и возможности, которые могут быть полезны при прогнозировании. Рассмотрим использованные модели.

Линейная модель является классической, и для краткости ее описание будет опущено.

Ridge-регрессия (*англ.* Ridge Regression), или гребневая регрессия, – это один из методов понижения размерности. Часто его применяют для борьбы с переизбыточностью данных, когда независимые переменные коррелируют друг с другом (т.е. имеет место мультиколлинеарность) [24].

Далее рассмотрим более мощные модели прогнозирования, такие как ансамблевые модели машинного обучения.

Bagging (*англ.* Bootstrap Aggregating) – ансамблевый метаалгоритм, предназначенный для улучшения стабильности и точности алгоритмов машинного обучения, используемых в задачах классификации и регрессии. Алгоритм также уменьшает дисперсию и помогает избежать переобучения.

Метод случайного леса (*англ.* Random Forest) – алгоритм машинного обучения, заключающийся в использовании ансамбля решающих деревьев. Алгоритм сочетает в себе две основные идеи: Bagging и метод случайных подпространств. Алгоритм применяется для задач классификации, регрессии и кластеризации. Основная идея заключается в использовании большого ансамбля решающих деревьев, каждое из которых само по себе дает очень невысокое качество модели, но за счет их большого количества результат получается качественным [25].

Следующий метод, градиентный бустинг (*англ.* Gradient Boosting), – метод машинного обучения, который используется для построения моделей, основанных на ансамбле деревьев решений. Он работает путем последовательного добавления деревьев к модели, каждое следующее дерево настраивается на ошибки предыдущего. Градиентный бустинг стремится минимизировать ошибку предсказания, используя градиентный спуск.

GradientBoostingRegressor является одним из примеров реализации градиентного бустинга в библиотеке scikit-learn для задач регрессии. Он обучает модель на основе ансамбля деревьев решений, минимизируя среднеквадратичную ошибку (*англ.* Mean Squared Error, MSE) между прогнозами и фактическими значениями.

В качестве параметров модели были определены:

- `n_estimators`: этот параметр определяет количество деревьев, которые будут использоваться в градиентном бустинге. В данном случае было выбрано 200 деревьев, что было обосновано результатом таких моделей, как `RandomForestRegressor`;
- `learning_rate`: этот параметр определяет вклад каждого дерева в окончательное предсказание модели. Чем меньше `learning_rate`, тем меньше влияние каждого дерева. В нашем случае было выбрано значение 0,1 с целью определения относительно небольшого вклада каждого дерева;
- `random_state`: устанавливается начальное значение для генератора случайных чисел с целью получения одинаковых результатов при повторном запуске модели.

Далее рассмотрим такую модель градиентного бустинга, как экстремальный градиентный бустинг (*англ.* `Extreme Gradient Boosting`, `XGBRegressor`).

`XGBRegressor` – это модель, основанная на библиотеке `XGBoost`. Она представляет собой улучшенную версию алгоритма градиентного бустинга, предназначенную для эффективного обучения моделей машинного обучения.

Отличие `XGBRegressor` от `GradientBoostingRegressor` заключается в нескольких аспектах:

- скорость и эффективность: `XGBoost` предлагает ряд оптимизаций, которые делают его более эффективным по сравнению с классическим градиентным бустингом. Это включает в себя параллельное обучение, оптимизированные структуры данных и возможность использования более сложных функций потерь.
- регуляризация: `XGBoost` предлагает встроенные методы регуляризации, такие как L1 (Lasso) и L2 (Ridge) регуляризация, что позволяет уменьшить переобучение модели.
- обработка пропущенных значений: `XGBoost` имеет встроенную обработку пропущенных значений, что упрощает работу с данными, содержащими пропуски.

В контексте прогнозирования доходов от услуг связи `XGBRegressor` может показывать лучшие результаты по нескольким причинам:

- `XGBoost` хорошо масштабируется для работы с большими наборами данных, что часто характерно для прогнозирования доходов от услуг связи;
- `XGBoost` способен эффективно работать с разнообразными типами признаков, включая категориальные, числовые и текстовые;
- благодаря оптимизациям и возможностям регуляризации `XGBoost` может достигать более высокой точности прогнозирования по сравнению с классическим градиентным бустингом.

Однако стоит отметить, что выбор между `XGBRegressor` и `GradientBoostingRegressor` зависит от конкретной задачи и набора данных, поэтому всегда важно проводить тщательное сравнение моделей перед принятием решения.

Данные были предварительно разделены на обучающую и тестовую выборки с использованием функции `train_test_split` из библиотеки `scikit-learn`. Соотношение обучающей и тестовой выборок равно 75% и 25% соответственно.

Также для оценки качества полученных моделей применен механизм перекрестной проверки (*англ.* `cross-validation`, CV), или кросс-фолдинг, – это метод оценки производительности модели машинного обучения, который посредством разбиения на несколько частей (фолдов) выборки и неоднократного обучения на разных сетах данных помогает оценить, насколько хорошо модель обобщает данные, не переобучаясь на конкретном наборе данных.

Недостатком моделей машинного обучения, как правило, является сложность интерпретации полученных результатов. Одним из методов, помогающих интерпретировать полученные результаты, выступает SHAP (*англ.* SHapley Additive exPlanations), представляющий собой метод интерпретации моделей машинного обучения и позволяющий оценить влияние каждого признака на прогнозы модели. Результаты анализа `SHAP summary_plot` позволяют наглядно увидеть влияние каждого признака на зависимую переменную (доходы от услуг связи) и их важность.

	Доходы_от_услуг_связи	Объем_информации	Индексы_тарифов_услуги_связи	Численность_работников_телекомм	Почтовые_переводы	Базовый_ИПЦ	Среднемесяч_зар_плата
count	2380.000000	2380.000000	2380.000000	2380.000000	2380.000000	2380.000000	2380.000000
mean	2.651258	98.812796	100.669130	2.383065	0.081299	101.493252	45379.919996
std	1.389965	75.845588	1.619499	0.878779	0.077090	1.796615	21390.899245
min	0.190687	0.309495	87.330000	0.426065	0.004279	98.750000	19674.090000
25%	2.037720	49.585350	100.000000	1.813065	0.044593	100.560000	31498.165000
50%	2.383098	81.142783	100.180000	2.288087	0.066829	100.980000	38781.250000
75%	2.876653	130.164365	100.820000	2.877390	0.089039	101.712500	50778.975000
max	13.306638	872.043954	129.040000	5.272566	0.924866	111.870000	158501.900000

Рис. 1. Описательные статистики рассматриваемых данных

Fig. 1. Descriptive statistics of the data

Источник: составлено авторами с использованием языка программирования Python

Этот вид графика позволяет увидеть важность каждого признака для всех предсказаний в общем; на оси  $X$  отображается SHAP-значение, показывающее влияние признака на предсказание; каждый признак представлен своим столбцом, и высота столбца показывает, насколько важен этот признак для модели.

### Результаты и обсуждение

На рис. 1 приведены результаты описательной статистики данных, помогающей понять основные характеристики и разброс значений каждого фактора.

Рассмотрим полученные результаты. Прежде всего следует обратить внимание на то, что средние значения факторов выше медианы, что свидетельствует о правосторонней асимметрии распределения. Соответственно, можно говорить о большом количестве выбросов в правой части.

На рис. 2 представлены гистограммы для каждой из рассматриваемых переменных (использована функция hist), подтверждающие наличие асимметрии.

На рис. 3 представлен бокс-плот, позволяющий выявить наличие выбросов, центральную тенденцию (медиану), разброс данных (размах) и наличие симметрии в распределении. Построим графики по каждому из регионов для показателя «Доходы от услуг связи» при помощи библиотеки Seaborn.

На графике бокс-плота по доходам от услуг связи по субъектам выделяются Москва, Санкт-Петербург, Магаданская область и Чукотский автономный округ. Это означает, что эти регионы имеют значительные выбросы или экстремально высокие значения доходов от услуг связи. Такие выбросы могут исказить общую картину и выделяться среди других регионов как особенно успешные или необычные по этому показателю. Дальнейший анализ данных и возможное исключение выбросов из рассмотрения могут помочь в более точном анализе и интерпретации результатов. Однако для моделей машинного обучения, основанных на деревьях обучения, этого, как правило, не требуется.

За исключением описанных выше субъектов с экстремально высокими значениями показателя, в остальных субъектах можно наблюдать следующее:

- Бокс-плоты большинства регионов находятся приблизительно на одном уровне: это означает, что медианное значение доходов от услуг связи для большинства регионов примерно одинаково.

- Весомое различие размахов для каждого региона. Размах бокс-плота отражает разницу между минимальным и максимальным значением доходов от услуг связи. Различия в размахе могут указывать на различия в источниках доходов, структуре рынка связи или других факторах, влияющих на доходы от услуг связи.

Последующим шагом будет построение корреляционной матрицы при помощи библиотеки Pandas (рис. 4).

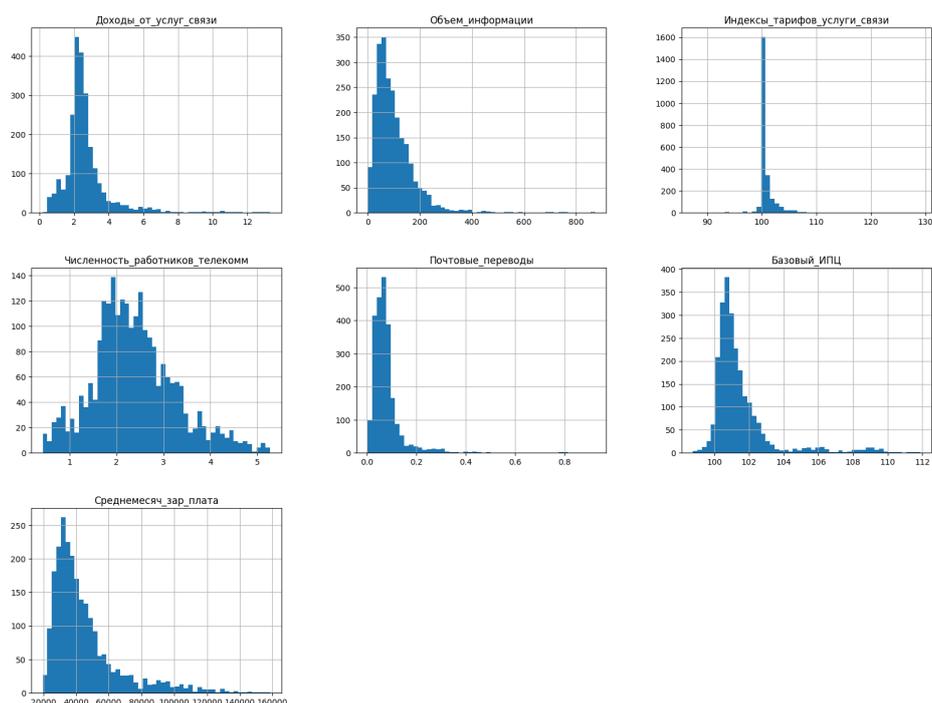


Рис. 2. Гистограммы для выбранных переменных

Fig. 2. Histograms for selected variables

Источник: составлено авторами с использованием языка программирования Python

В результате корреляционного анализа (рис. 4) можно увидеть, что есть достаточно сильно коррелирующие друг с другом переменные, например, такие как объем информации и среднемесячная заработная плата (0.47). Также есть факторы, практически не коррелирующие с другими переменными, например, такие как базовый индекс потребительских цен и индексы тарифов по услугам связи. Однако в целом можно сказать, что проблемы мультиколлинеарности не наблюдается.

Перейдем к построенным моделям.

Метрики качества всех моделей представлены в табл. 1.

Таблица 1. Сравнение результатов моделей прогнозирования

Table 1. Comparison of forecasting models' results

	Линейная регрессия	Ridge	Bagging	Random Forest Regressor	Gradient Boosting Regressor	XGB Regressor
$R^2$ для обучающей выборки	0,6182	0,6182	0,8821	0,881	0,7844	0,761
$R^2$ для тестовой выборки	0,5844	0,5843	0,7503	0,7527	0,9375	0,982
MSE для обучающей выборки	0,8472	0,8473	0,4691	0,469	0,4492	0,4979
MSE для тестовой выборки	0,9305	0,9306	0,3282	0,3282	0,1173	0,0339
$R$ -squared (средний)	0,6171	0,606	0,778	0,776	0,7954	0,7927

Источник: составлено авторами с использованием языка программирования Python

Из табл. 1 видно, что модель линейной регрессии, примененная к анализу доходов от услуг связи, показала следующие результаты:

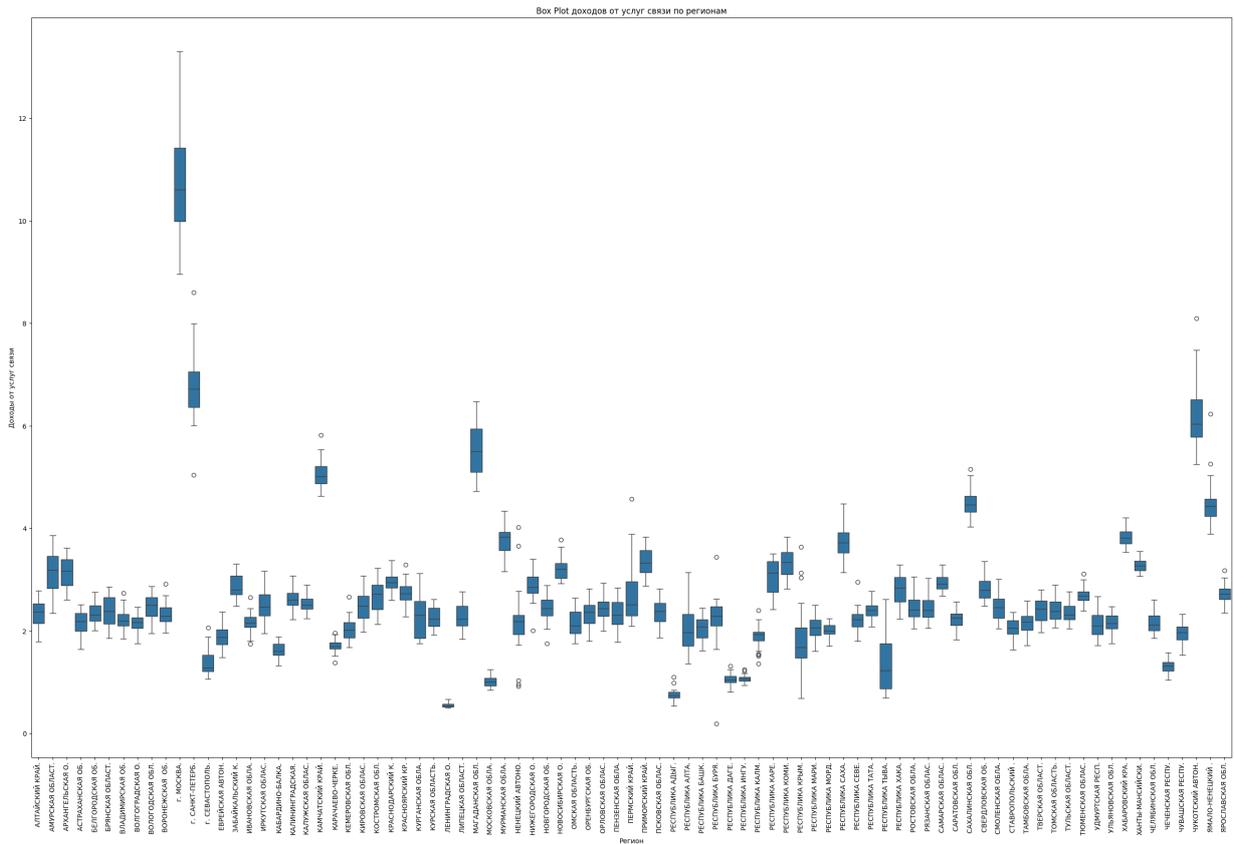


Рис. 3. График бокс-плотов по показателю доходов от услуг связи для всех регионов

Fig. 3. Boxplot graph of revenue of communication services for all regions

Источник: составлено авторами с использованием языка программирования Python

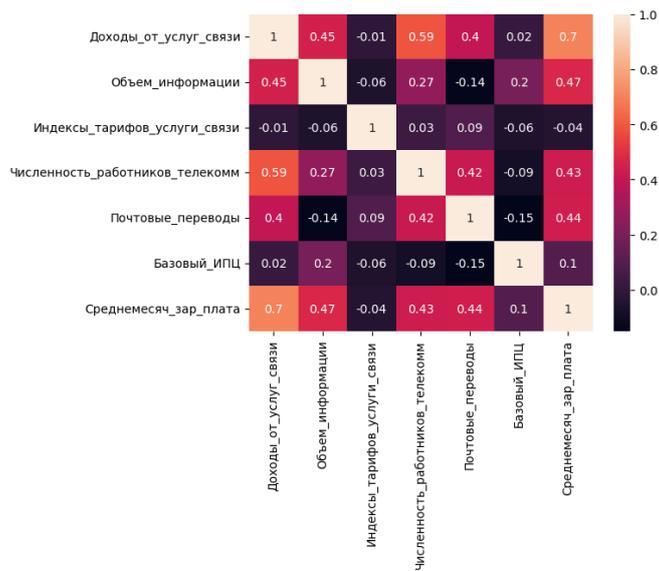


Рис. 4. Корреляционная матрица факторов

Fig. 4. Correlation matrix of factors

Источник: составлено авторами с использованием языка программирования Python

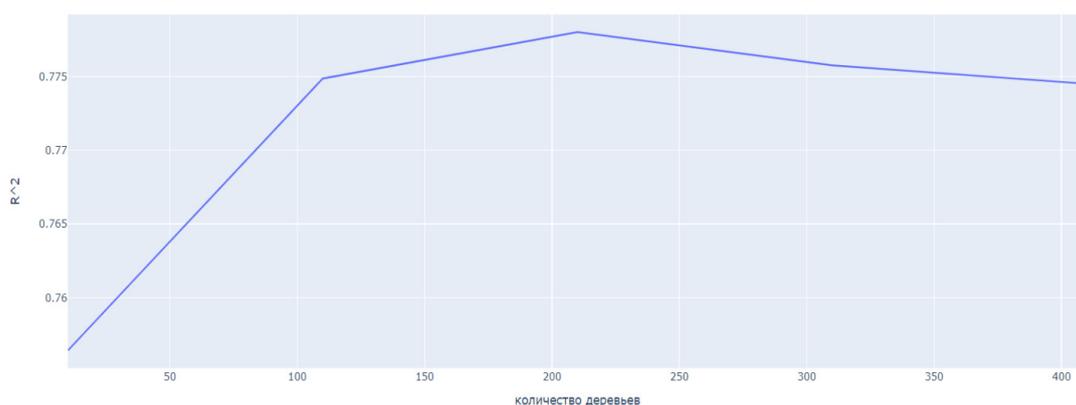


Рис. 5. Зависимость  $R^2$  от количества деревьев в модели Bagging

Fig. 5. Dependence of  $R^2$  on the number of trees in the Bagging model

Источник: составлено авторами с использованием языка программирования Python

–  $R^2$  для обучающей выборки – 0,6182;

–  $R^2$  для тестовой выборки – 0,5844.

Оценки коэффициента детерминации  $R^2$  для каждого из фолдов перекрестной проверки составили [0,629, 0,6861, 0,696, 0,6257, 0,6653, 0,5459, 0,5798, 0,5964, 0,5416, 0,6057]. Среднее  $R^2$  – 0,6171.

Среднеквадратичная ошибка при перекрестной проверке (CV MSE) – 0,8326.

Исходя из данных результатов, можно сделать вывод, что применение линейной регрессии для прогнозирования доходов от услуг связи показало среднюю точность и умеренную способность данной модели объяснять изменчивость целевой переменной. Для более точных прогнозов возможно рассмотрение других моделей.

Модель Ridge показала следующие результаты:

–  $R^2$  для обучающей выборки – 0,6182;

–  $R^2$  для тестовой выборки – 0,5843.

Коэффициенты  $R^2$ , полученные в модели Ridge, практически идентичны коэффициентам в модели линейной регрессии.

Модель, полученная при помощи Bagging, показала следующие результаты:

–  $R^2$  для обучающей выборки – 0,88;

–  $R^2$  для тестовой выборки – 0,75.

Кроме того, проведено исследование Bagging с разным количеством деревьев (10, 110, 210, 310, 410) в ансамбле, и результаты показали, что увеличение количества деревьев в ансамбле не всегда приводит к улучшению качества предсказаний на тестовой выборке (рис. 5), так как  $R^2$  на тестовой выборке достиг максимального значения при количестве деревьев равном 210.

RandomForestRegressor:

–  $R^2$  для обучающей выборки – 0,881;

–  $R^2$  для тестовой выборки – 0,753.

Кроме того, проведено аналогичное исследование с разным количеством деревьев в ансамбле (10, 110, 210, 310, 410), и результаты показали, что увеличение количества деревьев в ансамбле приводит к улучшению качества предсказаний на тестовой выборке, достигая наилучшего значения коэффициента детерминации при количестве деревьев равном 210 (рис. 6).

Среди моделей Ridge, Bagging и RandomForestRegressor наилучшие результаты детерминации показал метод RandomForestRegressor.

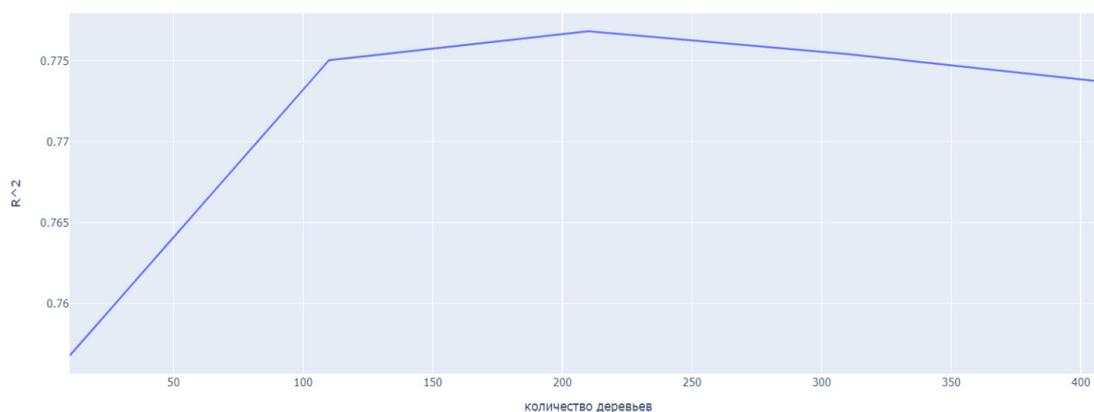


Рис. 6. Зависимость  $R^2$  от количества деревьев в модели RandomForestRegressor  
 Fig. 6. Dependence of  $R^2$  on the number of trees in the RandomForestRegressor model

Источник: составлено авторами с использованием языка программирования Python

Это объясняется следующими показателями:

- сложность: RandomForestRegressor является ансамблевым методом, который объединяет множество деревьев решений, позволяя каждому дереву вносить свой вклад в прогнозирование. Это позволяет модели лучше улавливать сложные нелинейные зависимости в данных и делать более точные прогнозы;
- устойчивость к переобучению: RandomForestRegressor обладает устойчивостью к переобучению благодаря случайному выбору подмножества признаков и объектов для построения каждого дерева, что позволяет улучшить обобщающую способность модели и предотвратить переобучение;
- эффективность в работе с мультиколлинеарностью: RandomForestRegressor является методом, который не чувствителен к сильной корреляции между признаками. Поэтому модель является более устойчивой к мультиколлинеарности, которая может влиять на другие линейные модели.

Таким образом, RandomForestRegressor показал наилучшие результаты детерминации благодаря своей способности моделировать сложные зависимости, устойчивости к переобучению и эффективности в работе с мультиколлинеарностью, что делает его мощным инструментом для прогнозирования.

Результаты GradientBoostingRegressor демонстрируют следующую производительность модели на тестовой и обучающей выборках:

- Значение среднеквадратичной ошибки прогнозов модели на тестовой выборке составляет 0,449, а на обучающей – 0,117. Чем ближе это значение к нулю, тем лучше.
- Значение коэффициента детерминации (англ.  $R^2$  score) на тестовой выборке составляет 0,784, а на обучающей – 0,938. Когда  $R^2$  для тестовой выборки значительно меньше, чем для обучающей, это может указывать на переобучение модели. Переобучение происходит, когда модель слишком хорошо подстраивается под обучающие данные и не может обобщить свои предсказания на новые данные. Здесь мы видим некоторое переобучение.

Для оценки производительности модели на нескольких различных разбиениях обучающей выборки возможно применение перекрестной проверки. В данном случае можно оценить модель с использованием 5-кратной перекрестной проверки. Коэффициент детерминации  $R^2$  для перекрестной проверки (англ. cross-validated  $R^2$  score) представляет собой массив значений  $R^2$ , полученных для каждого из 5 разбиений данных. Значения составляют 0,77219255; 0,73630886; 0,8424136; 0,75030561; 0,87604394.

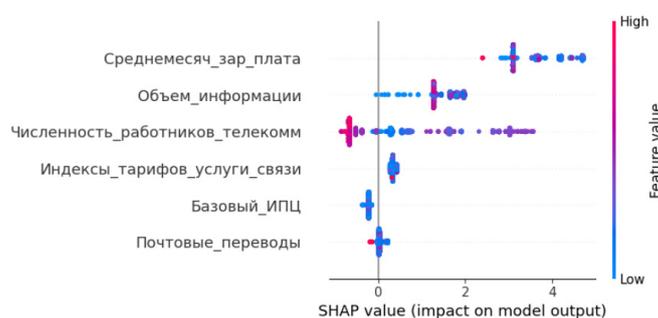


Рис. 7. Summary\_plot для модели GradientBoostingRegressor

Fig. 7. Summary\_plot for the GradientBoostingRegressor model

Источник: составлено авторами с использованием языка программирования Python

Среднее значение  $R^2$  по всем разбиениям составляет примерно 0,7954. Это означает, что в среднем модель объясняет около 79,54% дисперсии зависимой переменной в данных. Такие результаты указывают на то, что модель достаточно хорошо работает на нашем наборе данных и интерпретация влияния факторов на результирующий показатель, сделанная с ее помощью, имеет высокую степень надежности.

Эти результаты могут быть использованы для сравнения с другими моделями или для оценки улучшений в моделировании, таких как изменения параметров или использование других алгоритмов.

Результаты XGBRegressor демонстрируют следующую производительность модели на тестовой и обучающей выборках:

- Значение MSE для обучающей выборки (0,03) существенно ниже, чем для тестовой (0,49). Это может свидетельствовать о том, что модель хорошо справляется с предсказанием на обучающих данных, но менее точна на новых, тестовых. Однако MSE сама по себе не является определяющим показателем переобучения.

- Значение коэффициента детерминации указывает на возможное переобучение. Высокий  $R^2$  на обучающей выборке (0,982) говорит о том, что модель хорошо объясняет изменчивость данных в этом случае. Однако существенное снижение  $R^2$  на тестовой выборке (0,761) свидетельствует о переобучении.

Коэффициент детерминации  $R^2$  для перекрестной проверки – 0,73896302; 0,72659972; 0,8422616; 0,79155971; 0,86436171. Среднее значение  $R^2$  по всем разбиениям составляет примерно 0,7927.

Исходя из результатов приведенных выше моделей, можно предположить, что на нашем наборе данных более предпочтительной является модель GradientBoostingRegressor, у которой коэффициент детерминации для тестовой модели и при перекрестной проверке оказался выше. Однако для обучающей модели он выше, но все же выбор сделан в пользу GradientBoostingRegressor, поскольку разница между  $R^2$  для тестовой и обучающей выборок была менее значительной, чем у модели XGBRegressor.

Далее рассмотрим анализ SHAP-значений для модели GradientBoostingRegressor. В результате работы функции `summary_plot` был получен следующий график (рис. 7).

Рассмотрим подробно каждый фактор, представленный на графике:

#### 1. Среднемесячная заработная плата

SHAP-значения от 2,2 до 5 с модой значений в 3,1 указывают на то, что повышение среднемесячной заработной платы положительно влияет на доходы. Чем выше зарплата, тем вероятнее, что клиенты будут готовы тратить больше на услуги связи, что подтверждает гипотезу  $H_5$ .

#### 2. Объем информации



SHAP-значения от 0 до 2 с модой в 1,2 показывают, что увеличение объема информации также положительно влияет на доходы. Это может означать, что спрос на услуги связи растет с увеличением объема информации, который требуется передавать, что подтверждает гипотезу  $H_1$ .

### 3. Численность работников телекоммуникаций

SHAP-значения от  $-1$  до 3,7 с модой в  $-0,8$  указывают на то, что численность работников телекоммуникаций имеет разнонаправленное влияние на доходы от услуг связи, при этом размах SHAP-значений весьма велик. Отрицательные значения могут указывать на то, что увеличение численности работников может негативно сказываться на доходах, возможно, из-за дополнительных расходов на оплату персонала. Положительные значения могут указывать на улучшение качества обслуживания и, как следствие, на рост доходов, что отвергает гипотезу  $H_2$ .

### 4. Индексы тарифов услуг связи для юридических лиц

SHAP-значение около 0,3 указывает на то, что повышение индексов тарифов услуг связи положительно влияет на доходы. Это может означать, что клиенты готовы тратить больше денег на услуги связи, если тарифы увеличиваются. Возможно, это свидетельствует о высокой ценовой эластичности спроса в данной отрасли, что подтверждает гипотезу  $H_3$ .

### 5. Базовый индекс потребительских цен

SHAP-значение около  $-0,2$  показывает, что увеличение базового индекса потребительских цен может негативно влиять на доходы. Это может быть связано с тем, что при росте цен на потребительские товары и услуги клиенты могут быть более осторожными в расходовании денег на дополнительные услуги связи, что подтверждает гипотезу  $H_6$ .

### 6. Почтовые переводы

SHAP-значение от  $-0,3$  до 0,3 с модой в 0,1 указывает на то, что почтовые переводы имеют незначительное влияние на доходы. Этот фактор, скорее всего, не оказывает значительного влияния на спрос или доходы от услуг связи, однако может в некоторой степени подтверждать гипотезу  $H_4$ .

Интерпретация этих результатов позволяет лучше понять, какие факторы оказывают наибольшее влияние на доходы от услуг связи и какие изменения могут повлиять на результаты деятельности в данной отрасли.

Таким образом, `summary_plot` дает общее представление о важности признаков для модели. Рассмотрим также график `force_plot`, который позволяет анализировать влияние конкретных признаков на конкретные предсказания.

Основные признаки `force_plot`:

- этот вид графика позволяет увидеть влияние конкретных признаков на конкретное предсказание;
- возможен выбор конкретного предсказания для того, чтобы увидеть, какие признаки и в какую сторону влияют на прогнозное значение результирующего признака.

В нашем случае при выборе определенного региона за нужный квартал определенного года `force_plot` поможет увидеть, какие именно признаки и как влияют на предсказание для этого конкретного случая.

Рассмотрим на примере `force_plot` для Москвы на IV квартал 2023 г. (рис. 8).

На основании предоставленных данных можно сделать следующие выводы о ситуации в Москве в IV квартале 2023 г.:

- Среднемесячная зарплата имеет наибольшее положительное влияние на доходы от услуг связи, что говорит о том, что увеличение среднемесячной зарплаты приводит к их увеличению.
- Объем информации также оказывает значительное положительное влияние на доходы от услуг связи. Увеличение объема информации способствует их увеличению.
- Индексы тарифов на услуги связи и почтовые переводы имеют меньшее, но все же положительное влияние на доходы от них.

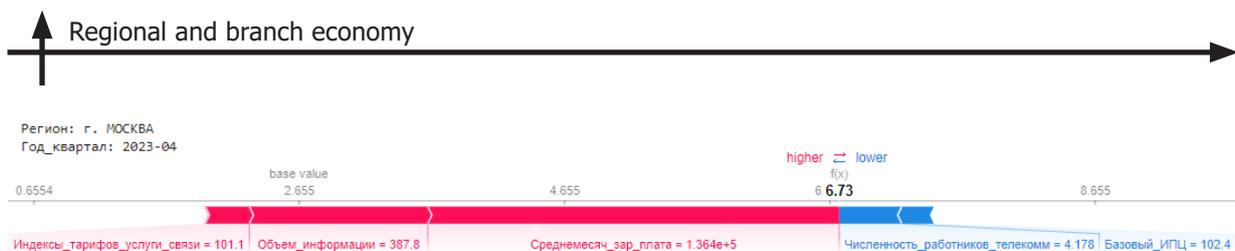


Рис. 8. Force\_plot для модели GradientBoostingRegressor на IV квартал 2023 г. для Москвы

Fig. 8. Force\_plot for the GradientBoostingRegressor model for Q4 2023 for Moscow

Источник: составлено авторами с использованием языка программирования Python

— Численность работников в отрасли телекоммуникаций и базовый индекс потребительских цен оказывают негативное влияние на доходы от услуг связи.

Исходя из всего вышесказанного, можно сделать вывод, что SHAP-значения и GradientBoostingRegressor могут быть полезны для анализа сектора связи и телекоммуникаций. Они позволяют выявить ключевые факторы, влияющие на доходы от услуг связи, и строить модели, способные учитывать сложные взаимосвязи между различными параметрами. GradientBoostingRegressor является полезным инструментом для прогнозирования в условиях наличия независимых переменных. Его применение на практике оправдано, особенно при необходимости построения сложных моделей для прогнозирования. Таким образом, эти модели могут быть использованы для достаточно точного прогнозирования и оптимизации бизнес-процессов в секторе связи и телекоммуникаций.

### Заключение

В рамках данного исследования в соответствии с поставленными задачами получены следующие результаты:

1. Рассмотрены основные примеры работ из отечественной и зарубежной научной литературы, посвященных теме развития сектора телекоммуникаций и связи. Сделан вывод о том, что факторов, оказывающих на него воздействие, значительное количество. Однако обзор литературы показал, что не учитывается региональная специфика. Восполнение данного пробела в научных исследованиях и стало целью исследования.

2. Проведены сбор и систематизация статистических данных для регионов Российской Федерации за 2017–2023 гг. (поквартально). В результате количество наблюдений в рамках панели составило 2380, что позволяет использовать достаточно сложные методы статистического анализа и машинного обучения, включая линейную регрессию, Ridge, Bagging, RandomForestRegressor, GradientBoostingRegressor и XGBRegressor. В качестве результирующей переменной рассматриваются доходы от услуг связи.

3. Осуществлена описательная статистика данных, анализ выбросов, корреляционный анализ. Так, сделаны выводы о наличии правосторонней асимметрии распределения, что позволяет говорить о большом количестве выбросов в правой части. Из субъектов России выделяются Москва, Санкт-Петербург, Магаданская область и Чукотский автономный округ. Это означает, что упомянутые регионы имеют значительные выбросы или экстремально высокие значения доходов от услуг связи по сравнению с другими. Такие выбросы могут исказить общую картину и при использовании ряда классических методов желательное исключение выбросов из рассмотрения, однако для моделей машинного обучения, основанных на деревьях обучения, этого, как правило, не требуется.

4. Модель GradientBoostingRegressor показала лучшие результаты при анализе рассматриваемых данных. С целью интерпретации полученных результатов проанализированы SHAP-значения, которые помогли выявить характер влияния различных факторов на развитие сектора связи и телекоммуникаций, таких как повышение среднемесячной заработной платы, увеличение объема информации



(трафика), индексы тарифов услуг связи для юридических лиц, а также негативное влияние на доходы от услуг связи базового индекса потребительских цен и при этом незначительное влияние на них почтовых переводов, которые в последующем при формировании подобных моделей можно не учитывать.

Итак, полученные результаты анализа сектора связи и телекоммуникаций могут быть использованы для различных целей:

– Стратегическое планирование. Результаты анализа могут помочь в формировании стратегий развития сектора связи и телекоммуникаций на уровне регионов. Это может включать определение приоритетных направлений развития, выявление регионов с наибольшим потенциалом для роста доходов от услуг связи и разработку целевых планов действий.

– Оптимизация инвестиций. Полученные данные могут помочь инвесторам и компаниям в принятии решений об инвестировании в сектор связи и телекоммуникаций. На основе анализа можно определить регионы с наибольшим потенциалом для роста и оптимизировать распределение инвестиций.

– Государственное регулирование. Результаты анализа могут быть использованы для разработки государственной политики в области связи и телекоммуникаций. Например, они могут помочь в определении регионов, требующих дополнительной поддержки или стимулирования развития данного сектора.

– Маркетинг и продажи. Для компаний в секторе связи и телекоммуникаций результаты анализа могут быть использованы в определении потенциальных рынков для продвижения продуктов и услуг, а также для уточнения маркетинговых стратегий в зависимости от специфики каждого региона.

– Прогнозирование и планирование: Полученные данные могут быть использованы для прогнозирования динамики доходов от услуг связи в различных регионах, что поможет компаниям и государственным органам более точно планировать свои действия и ресурсы.

Таким образом, результаты анализа могут оказать значительное влияние на различные аспекты управления и развития сектора связи и телекоммуникаций, помогая принимать обоснованные решения и оптимизировать деятельность как на уровне компаний, так и на уровне государственной политики.

#### **Направления дальнейших исследований**

Несмотря на то, что качество модели с учетом количества рассмотренных наблюдений представляется вполне высоким, в полученных моделях можно видеть некоторое переобучение. Вероятнее всего, его уровень был бы ниже, если бы выборка была больше. По мере появления данных за новые периоды целесообразным представляется увеличение наблюдений для уточнения параметров модели. Кроме того, целесообразным представляется по мере увеличения выборки генерировать нейронные сети для получения еще более точных оценок влияния различных факторов на сектор связи и телекоммуникаций, а также для получения более точных прогнозных оценок.

#### **СПИСОК ИСТОЧНИКОВ**

1. Oumbé H.T., Nougan R.D., Ndzana A.M. (2023) Does information and communication technologies affect economic complexity? *SN Business & Economics*, 3, art. no. 92. DOI: <https://doi.org/10.1007/s43546-023-00467-8>

2. Gómez-Barroso J.L., Marbán-Flores R. (2020). Telecommunications and economic development – The 21st century: Making the evidence stronger. *Telecommunications Policy*, 44 (2), art. no. 101905. DOI: <https://doi.org/10.1016/j.telpol.2019.101905>

3. Latif Z., Yang M., Danish, Ximei L., Pathan Z., Shafaq S., Jianqiu Z. (2017) The dynamics of ICT, foreign direct investment, globalization and economic growth: Panel estimation robust to hetero-

geneity and cross-sectional dependence. *Telematics and Informatics*, 35 (2), 318–328. DOI: <https://doi.org/10.1016/j.tele.2017.12.006>

4. Coccia M. (2020) Asymmetry of the technological cycle of disruptive innovations. *Technology Analysis & Strategic Management*, 32 (12), 1462–1477. DOI: <https://doi.org/10.1080/09537325.2020.1785415>

5. Hitt L. (1999) Information Technology and Firm Boundaries: Evidence from Panel Data. *Information Systems Research*, 10 (2), 134–149. DOI: <https://doi.org/10.1287/isre.10.2.134>

6. Gholami R., Lee S.-Y., Heshmati A. (2006) The Causal Relationship Between Information and Communication Technology and Foreign Direct Investment. *World Economy*, 29 (1), 43–62. DOI: <https://doi.org/10.1111/j.1467-9701.2006.00757.x>

7. Фокина Е.А. (2014) Телекоммуникационный сектор экономики России: трансформация в глобальную информационно-телекоммуникационную инфраструктуру. *Вестник Волгоградского государственного университета. Серия 3: Экономика. Экология*, 6 (29), 103–113. DOI: <http://dx.doi.org/10.15688/jvolsu3.2014.6.11>

8. Левичев И.В., Васильев С.А., Канзитдинов Ш.К., Перес А.Д. (2024) Численный анализ модели динамического ценообразования на услуги 5G/6G для случая олигополии на рынке телекоммуникаций. *International Journal of Open Information Technologies*, 12 (1), 92–100.

9. Нусипова Ж.Е. (2019) Исследование тенденций, закономерностей, условий развития сектора телекоммуникаций в системе устойчивого функционирования экономики регионов. *Петербургский экономический журнал*, 1, 89–96. DOI: <https://doi.org/10.25631/PEJ.2019.1.89.96>

10. Школьник И.С. (2019) Институциональные условия диффузии инноваций в области телекоммуникаций: на примере Китая. *Инновации и инвестиции*, 1, 25–28.

11. Хатунцева Е.А., Хатунцев А.Б. (2016) Анализ основных тенденций развития сетей связи на телекоммуникационном рынке России. *T-Comm: Телекоммуникации и транспорт*, 10 (7), 71–74.

12. Хатунцева Е.А., Хатунцев А.Б. (2017) Некоторые аспекты дальнейшего совершенствования сетей связи в РФ. *Экономика и качество систем связи*, 1 (3), 53–59.

13. Бородин А.С., Волков А.Н., Мутханна А.С., Кучерявый А.Е. (2021) Искусственный интеллект в сетях связи пятого и последующих поколений. *Электросвязь*, 1, 17–22.

14. Katz R., Jung J. (2023) The impact of taxation in the telecommunications industry. *Information Economics and Policy*, 62, art. no. 101016. DOI: <https://doi.org/10.1016/j.infoecopol.2023.101016>

15. Yilanci V., Kilci E.N. (2024) A study on the macroeconomic and financial determinants of telecommunication infrastructure: Evidence from Turkiye. *Technology in Society*, 77, art. no. 102573. DOI: <https://doi.org/10.1016/j.techsoc.2024.102573>

16. Улезлова Л.В. (2017) Структура, тенденции и факторы развития рынка отрасли телекоммуникаций и связи. *Проблемы современной науки и образования*, 13 (95), 76–79. DOI: <https://doi.org/10.20861/2304-2338-2017-95>

17. Малкина М.Ю. (2023) Финансовое заражение отраслей российской экономики от нефтяных шоков в период пандемии. *Terra Economicus*, 21 (2), 6–22. DOI: <https://doi.org/10.18522/2073-6606-2023-21-2-6-22>

18. Галимов И.Р. (2021) Влияние телекоммуникаций на экономическое развитие регионов в западной и восточной части России. *Вестник Алтайской академии экономики и права*, 11 (2), 165–171. DOI: <https://doi.org/10.17513/vaael.1931>

19. Суглобова К.Ю. (2020) Экономический эффект от цифровизации отраслей реального сектора экономики в России. Требования к сетям связи нового поколения. *REDS: Телекоммуникационные устройства и системы*, 10 (4), 30–38.

20. Тонких Е.В. (2022) Анализ характеристик плотности устройств в сетях связи пятого поколения. *Информационные технологии и телекоммуникации*, 8 (1), 22–27. DOI: <https://doi.org/10.31854/2307-1303-2020-8-1-22-27>

21. Дьяков А.М., Фишер А.В. (2019) Состояние и перспективы развития телекоммуникаций в контексте формирования цифровой экономики в РФ. *Экономика и бизнес: теория и практика*, 6–1 (52), 120–122. DOI: <https://doi.org/10.24411/2411-0450-2019-10833>

22. Завьялов А.В., Поддубнов И.В. (2018) Эволюция сетей сотовой связи. *Экономика и социум*, 1 (44), 1372–1375.

23. Лазарян А.В. (2024) Новые методы повышения эффективности бизнеса операторов связи. *Инновации и инвестиции*, 4, 167–170.

24. Бикетов А.Р., Бею В.В., Свиридов И.Э. (2023) Анализ прикладных инструментов для разработки интеллектуальной системы контроля на этапе розлива ликера. *Фабрика будущего*:



переход к передовым цифровым, интеллектуальным производственным технологиям, роботизированным системам для отраслей пищевой промышленности: сборник научных докладов IV Международной специализированной конференции-выставки (Москва–2023), 20–26.

25. Бадыкова И.Р. (2024) Выявление факторов эффективности бизнеса: моделирование методом случайного леса. *Экономический анализ: теория и практика*, 23 (7), 1225–1239. DOI: <https://doi.org/10.24891/ea.23.7.1225>

## REFERENCES

1. Oumbé H.T., Nougan R.D., Ndzana A.M. (2023) Does information and communication technologies affect economic complexity? *SN Business & Economics*, 3, art. no. 92. DOI: <https://doi.org/10.1007/s43546-023-00467-8>

2. Gómez-Barroso J.L., Marbán-Flores R. (2020). Telecommunications and economic development – The 21st century: Making the evidence stronger. *Telecommunications Policy*, 44 (2), art. no. 101905. DOI: <https://doi.org/10.1016/j.telpol.2019.101905>

3. Latif Z., Yang M., Danish, Ximei L., Pathan Z., Shafaq S., Jianqiu Z. (2017) The dynamics of ICT, foreign direct investment, globalization and economic growth: Panel estimation robust to heterogeneity and cross-sectional dependence. *Telematics and Informatics*, 35 (2), 318–328. DOI: <https://doi.org/10.1016/j.tele.2017.12.006>

4. Coccia M. (2020) Asymmetry of the technological cycle of disruptive innovations. *Technology Analysis & Strategic Management*, 32 (12), 1462–1477. DOI: <https://doi.org/10.1080/09537325.2020.1785415>

5. Hitt L. (1999) Information Technology and Firm Boundaries: Evidence from Panel Data. *Information Systems Research*, 10 (2), 134–149. DOI: <https://doi.org/10.1287/isre.10.2.134>

6. Gholami R., Lee S.-Y., Heshmati A. (2006) The Causal Relationship Between Information and Communication Technology and Foreign Direct Investment. *World Economy*, 29 (1), 43–62. DOI: <https://doi.org/10.1111/j.1467-9701.2006.00757.x>

7. Fokina E.A. (2014) Telecommunication sector of the Russian economy: transformation into a global information and telecommunication infrastructure. *Science journal of Volgograd state university. Global economic system*, 6 (29), 103–113. DOI: <http://dx.doi.org/10.15688/jvolsu3.2014.6.11>

8. Levichev I.V., Vasilyev S.A., Kanzitdinov S.K., Peres Acosta D. (2024) Numerical analysis of the dynamic pricing model for 5G/6G services for oligopoly telecommunications markets. *International Journal of Open Information Technologies*, 12 (1), 92–100.

9. Nusipova Zh. E. (2019) The study of trends, patterns, conditions for the telecommunications sector development in the sustainable system of regional economy. *St. Petersburg Economic Journal*, 1, 89–96. DOI: <https://doi.org/10.25631/PEJ.2019.1.89.96>

10. Shkol'nik I.S. (2019) Institutional conditions of diffusion innovation in the field of telecommunications: the example of China. *Innovations and Investments*, 1, 25–28.

11. Khatuntseva E.A. Khatuntsev A.B. (2016) Analysis of major trends of communication networks development in Russia. *T-Comm*, 10 (7), 71–74.

12. Khatuntseva E.A. Khatuntsev A.B. (2017) Some aspects of further improvement of communication network in Russia. *Ekonomika i kachestvo sistem svyazi [Economy and quality of communication systems]*, 1 (3), 53–59.

13. Borodin A.S., Volkov A.N., Mutkhanna A.S., Kucheriavyi A.E. (2021) Iskusstvennyi intellekt v setiakh svyazi piatogo i posleduiushchikh pokolenii [Artificial Intelligence in Fifth and Subsequent Generation Communication Networks]. *Electrosvyaz Magazine*, 1, 17–22.

14. Katz R., Jung J. (2023) The impact of taxation in the telecommunications industry. *Information Economics and Policy*, 62, art. no. 101016. DOI: <https://doi.org/10.1016/j.infoecopol.2023.101016>

15. Yilanci V., Kilci E.N. (2024) A study on the macroeconomic and financial determinants of telecommunication infrastructure: Evidence from Türkiye. *Technology in Society*, 77, art. no. 102573. DOI: <https://doi.org/10.1016/j.techsoc.2024.102573>

16. Ulezlova L.V. (2017) Structure, tendencies and factors of development of the market in branch the telecommunications and communications. *Problems of modern science and education*, 13 (95), 76–79. DOI: <https://doi.org/10.20861/2304-2338-2017-95>

17. Malkina M.Yu. (2023) Financial contagion from oil shocks during the pandemic: A cross-sector analysis. *Terra Economicus*, 21 (2), 6–22. DOI: <https://doi.org/10.18522/2073-6606-2023-21-2-6-22>
18. Galimov I.R. (2021) Influence of telecommunications on economic development of regions in Western and Eastern Russia. *Vestnik Altaiskoi akademii ekonomiki i prava [Bulletin of the Altai Academy of Economics and Law]*, 11 (2), 165–171. DOI: <https://doi.org/10.17513/vaael.1931>
19. Suglobova K.IU. (2020) Ekonomicheskii effekt ot tsifrovizatsii otraslei real'nogo sektora ekonomiki v Rossii. Trebovaniia k setiam sviazi novogo pokoleniia [Economic effect from digitalization of real economy sectors in Russia. Requirements for new generation communication networks]. *REDS: Telekommunikatsionnye ustroistva i sistemy [REDS: Telecommunication devices and systems]*, 10 (4), 30–38.
20. Tonkikh E. (2020) The dense of devices analysis for 5G networks. *Telecom IT*, 8 (1), 22–27. DOI: <https://doi.org/10.31854/2307-1303-2020-8-1-22-27>
21. Dyakov A.M., Fisher A.V. (2019) Condition and perspectives of the development of telecommunications in the context of forming a digital economy in the Russian Federation. *Journal of Economy and Business*, 6–1 (52), 120–122. DOI: <https://doi.org/10.24411/2411-0450-2019-10833>
22. Zavyalov A.V., Poddubnov I.V. (2018) Evolution of cellular networks. *Ekonomika i sotsium [Economy and society]*, 1 (44), 1372–1375.
23. Lazaryan A.V. (2024) New methods for enhancing the efficiency of telecommunication operators. *Innovations and Investments*, 4, 167–170.
24. Biketov A.R., Beiu V.V., Sviridov I.E. (2023) Analiz prikladnykh instrumentov dlia razrabotki intell'ektual'noi sistemy kontroliia na etape rozliva likera [Analysis of applied tools for developing intelligent control system at the stage of liquor bottling]. *Fabrika budushchego: perekhod k peredovym tsifrovym, intellektual'nym proizvodstvennym tekhnologiiam, robotizirovannym sistemam dlia otraslei pishchevoi promyshlennosti: sbornik nauchnykh dokladov IV Mezhdunarodnoi spetsializirovannoi konferentsii-vystavki (Moskva–2023) [Factory of the Future: Transition to Advanced Digital, Intelligent Manufacturing Technologies, Robotic Systems for the Food Industry: Collection of Scientific Papers of the IV International Specialized Conference and Exhibition (Moscow–2023)]*, 20–26.
25. Badykova I.R. (2024) Identifying business efficiency factors: Random forest modeling. *Economic Analysis: Theory and Practice*, 23 (7), 1225–1239. DOI: <https://doi.org/10.24891/ea.23.7.1225>

## СВЕДЕНИЯ ОБ АВТОРАХ / INFORMATION ABOUT AUTHORS

**БАДЫКОВА Иделя Рашитовна**

E-mail: [idelia.badykova@gmail.com](mailto:idelia.badykova@gmail.com)

**Idelya R. BADYKOVA**

E-mail: [idelia.badykova@gmail.com](mailto:idelia.badykova@gmail.com)

ORCID: <https://orcid.org/0000-0002-9072-3856>

**БИКТИМИРОВА Камилла Романовна**

E-mail: [biktimirovakamilla0@gmail.com](mailto:biktimirovakamilla0@gmail.com)

**Kamilla R. BIKTIMIROVA**

E-mail: [biktimirovakamilla0@gmail.com](mailto:biktimirovakamilla0@gmail.com)

*Поступила: 29.09.2024; Одобрена: 03.12.2024; Принята: 03.12.2024.*

*Submitted: 29.09.2024; Approved: 03.12.2024; Accepted: 03.12.2024.*