

УДК 004.89

Научная статья

DOI: 10.35330/1991-6639-2025-27-1-143-151

EDN: XRYMDH

## Сравнительный анализ методов снижения дисбаланса классов при построении моделей машинного обучения в финансовом секторе

А. Ф. Константинов<sup>✉</sup>, Л. П. Дьяконова

Российский экономический университет имени Г. В. Плеханова  
115054, Россия, Москва, Стремянный переулок, 36

**Аннотация.** В статье рассматриваются методы улучшения показателей качества моделей машинного обучения, применяемых в финансовом секторе. В связи с тем, что наборы данных, на которых обучаются модели, обладают несбалансированностью классов, предлагается использовать модели, направленные на снижение дисбаланса. В исследовании были проведены эксперименты с применением 9 методов учета несбалансированности классов к трем наборам данных по розничному кредитованию. В качестве базовой использовалась модель градиентного бустинга CatboostClassifier, не учитывающая дисбаланс классов. Проведенные эксперименты показали, что применение метода RandomOverSampler дает существенный прирост показателей качества классификации по сравнению с базовой моделью. Результаты свидетельствуют о перспективности дальнейших исследований методов учета дисбаланса классов при изучении финансовых данных, а также о целесообразности применения рассмотренных методов на практике.

**Ключевые слова:** финансовые риски, машинное обучение, классификация, дисбаланс классов

Поступила 24.09.2024, одобрена после рецензирования 25.10.2024, принята к публикации 09.01.2025

**Для цитирования.** Константинов А. Ф., Дьяконова Л. П. Сравнительный анализ методов снижения дисбаланса классов при построении моделей машинного обучения в финансовом секторе // Известия Кабардино-Балкарского научного центра РАН. 2025. Т. 27. № 1. С. 143–151. DOI: 10.35330/1991-6639-2025-27-1-143-151

MSC: 90C99

Original article

## Comparative analysis of class imbalance reduction methods in building machine learning models in the financial sector

A.F. Konstantinov<sup>✉</sup>, L.P. Dyakonova

Plekhanov Russian University of Economics  
115054, Russia, Moscow, 36 Stremyanny lane

**Abstract.** The article discusses methods for improving quality metrics of machine learning models used in the financial sector. Due to the fact that the data sets on which the models are trained have class imbalances, it is proposed to use models aimed at reducing the imbalance. The study conducted experiments using 9 methods for accounting for class imbalances with three data sets on retail lending. The CatboostClassifier gradient boosting model, which does not take into account class imbalances, was used as the base model. The experiments showed that the use of the RandomOverSampler method provides a significant increase in classification quality metrics compared to the base model. The results indicate the

promise of further research into methods for accounting for class imbalances in the study of financial data, as well as the feasibility of application of the considered methods in practice.

**Keywords:** financial risks, machine learning, classification, class imbalance

Submitted 24.09.2024,

approved after reviewing 25.10.2024,

accepted for publication 09.01.2025

**For citation.** Konstantinov A.F., Dyakonova L.P. Comparative analysis of class imbalance reduction methods in building machine learning models in financial sector. *News of the Kabardino-Balkarian Scientific Center of RAS.* 2025. Vol. 27. No. 1. Pp. 143–151. DOI: 10.35330/1991-6639-2025-27-1-143-151

## ВВЕДЕНИЕ

Деятельность компаний финансового сектора связана с высокой степенью риска. Для коммерческих банков к основным рискам относятся кредитный и операционный. На финансовом рынке кредитование сохраняет позицию наиболее доходной статьи активов кредитных организаций, хотя и наиболее рискованной. Кредитный риск, таким образом, был и остается основным видом банковского риска.

По данным Центрального банка Российской Федерации, в настоящее время наблюдается значительный рост кредитных дефолтов и мошеннических операций. На конец 2 квартала 2024 г. число заемщиков банков и МФО превысило 50 млн человек, причем за год число заемщиков выросло на 3,6 млн человек, а сумма задолженности возросла на 7,7 трлн рублей<sup>1</sup>. В 2023 году объем операций без согласия клиентов увеличился по сравнению с 2022-м на 11,48 %<sup>2</sup>.

Начиная с 2023 года кредитные организации направляют в Банк России информацию о предотвращенных хищениях денежных средств. Так, за 2023 год объем предотвращенных операций без согласия клиентов составил 5 798,35 млрд рублей. Благодаря эффективной работе антифрод-процедур кредитных организаций злоумышленникам не удалось совершить 34,77 млн мошеннических операций.

В настоящее время финансовыми организациями для минимизации указанных рисков широко используются процедуры, основанные на методах машинного обучения и искусственного интеллекта, в которых проводится обучение моделей на основании исторических данных. Характерной особенностью таких наборов данных является значительный дисбаланс классов, то есть частота положительного класса (дефолта, мошенничества) значительно меньше частоты отрицательного класса (нормальные операции клиентов).

Современные алгоритмы могут работать с данными, имеющими высокий дисбаланс классов. В то же время существует целый ряд методов работы с дисбалансом в данных, которые позволяют дополнительно улучшить показатели качества моделей машинного обучения. В настоящий момент для снижения дисбаланса известны<sup>3</sup>:

- 8 методов, увеличивающих количество наблюдений положительного класса;
- 11 методов, снижающих количество наблюдений отрицательного класса;

<sup>1</sup>Отчет Центрального Банка РФ. Обзор операций, совершенных без согласия клиентов финансовых организаций. [https://www.cbr.ru/analytics/ib/operations\\_survey/2023/](https://www.cbr.ru/analytics/ib/operations_survey/2023/) (дата обращения: 26.11.2024).

<sup>2</sup>Отчет Центрального Банка РФ. Анализ тенденций в сегменте розничного кредитования на основе данных бюро кредитных историй. Первое полугодие 2024 года. [https://cbr.ru/Collection/Collection/File/50679/inf-material\\_bki\\_2024fh.pdf](https://cbr.ru/Collection/Collection/File/50679/inf-material_bki_2024fh.pdf) (дата обращения: 26.11.2024).

<sup>3</sup>Imbalanced Learn. API reference. available at: <https://imbalanced-learn.org/stable/references/index.html> (дата обращения: 26.11.2024).

- 2 комбинированных метода.

Однако далеко не все эти методы демонстрируют улучшение показателей качества классификации на финансовых данных. Настоящее исследование направлено на определение методов снижения дисбаланса классов, которые наилучшим образом подходят для применения при обучении моделей машинного обучения в финансовом секторе. Результаты исследования позволят определить наиболее эффективные методы и модели для работы с финансовыми данными.

**Цели и задачи исследования:** провести сравнение методов снижения дисбаланса классов при обучении моделей машинного обучения на финансовых данных и определить методы, которые дают повышение показателей качества классификации на финансовых данных.

#### ОПИСАНИЕ НАБОРА ДАННЫХ

**Таблица 1.** Наборы данных для экспериментов

**Table 1.** Datasets for the experiments

№	Наименование	Отрасль	Количество строк	Количество столбцов	Доля положительного класса
1	Loan Defaulter	Финансы	307 511	122	0.08
2	Loan Default Prediction Dataset	Финансы	255 347	18	0.12
3	Bank Account Fraud Dataset Suite (NeurIPS 2022)	Финансы	1 000 000	32	0.01

##### 1. Loan Defaulter<sup>4</sup>.

Набор данных содержит информацию для исследовательского анализа данных в реальном бизнес-сценарии в банковской сфере для минимизации риска потерь при кредитовании клиентов. Набор данных содержит 122 переменные.

Классы информации в наборе данных:

- информация о клиенте (возраст, пол, сведения о работе, информация о месте жительства клиента и т.д.);
- информация о внешних скорингах;
- сведения о предоставленных документах.

Целевая переменная: был ли дефолт по этому кредиту (0/1).

##### 2. Loan Default Prediction Dataset<sup>5</sup>.

Набор данных содержит 255 347 строк и 18 переменных.

Классы информации в наборе данных:

- информация о заемщике (пол, возраст, образование, тип работы, статус состояния в браке и т.д.);
- информация об истории кредитования (кредитный балл, количество месяцев кредита, количество кредитов, процентная ставка и т.д.).

Целевая переменная: был ли дефолт по этому кредиту (0/1).

##### 3. Bank Account Fraud Dataset Suite<sup>6</sup>.

<sup>4</sup>Ссылка: <https://www.kaggle.com/datasets/gauravduttakiit/loan-defaulter>

<sup>5</sup>Ссылка: <https://www.kaggle.com/datasets/nikhil1e9/loan-default>

<sup>6</sup>Ссылка: <https://www.kaggle.com/datasets/sgpjesus/bank-account-fraud-dataset-neurips-2022>

Набор данных Bank Account Fraud (BAF) был опубликован на NeurIPS 2022. Для целей исследования брался только базовый вариант данных. Для защиты личности потенциальных заявителей авторами набора данных были применены дифференциальные методы конфиденциальности (добавление шума), кодирование признаков.

Классы информации в наборе данных:

- информация о заявителе (доход, возраст, статус работы, статус наличия жилья, количество месяцев обслуживания банком и т.д.);
- информация о запросе (время от запроса, тип платежа, зарубежный запрос, операционная система гаджета и т.д.);
- внешние скоринги;
- логические противоречия (телефон домашний/мобильный валидный, не сгенерирован ли email автоматически и т.д.).

Целевая переменная: был ли дефолт по этому кредиту (0/1).

#### СХЕМА ВАЛИДАЦИИ

Для каждого набора данных наблюдения разделялись методом кросс-валидации StratifiedShuffleSplit<sup>7</sup> на тренировочные (50 %) и тестовые (50 %) данные, и для каждого набора тренировочных данных 3 раза проводилось разделение на тренировочный и валидационный набор данных с помощью метода кросс-валидации kFold<sup>8</sup> (20% на валидационные данные, 80 % – данные для обучения).

#### ОПИСАНИЕ БАЗОВОЙ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ

В качестве базовой модели для моделирования использовалась модель градиентного бустинга CatboostClassifier, разработанная Яндекс. Модель представляет ансамбль деревьев решений небольшой глубины, причем на каждой последующей итерации модель учится снижать псевдоостатки прогнозов предыдущих итераций деревьев. При обучении были установлены следующие гиперпараметры модели: 'iterations'=3000, 'early\_stopping\_rounds' = 100, 'eval\_set' = (X\_val, y\_val). После разбиения каждого набора данных на сплит и фолд на данных для обучения был обучен кодировщик категориальных переменных TargetEncoder<sup>9</sup> из библиотеки category\_encoders. Так как при обучении TargetEncoder мы использовали только данные для обучения, утечки данных не происходит. Далее были преобразованы категориальные переменные в валидационных и тестовых наборах данных X\_val, X\_test. Применение отдельного кодировщика, а не встроен-

<sup>7</sup>Метод кросс-валидации StratifiedShuffleSplit – это метод, используемый для разделения данных на тренировочные и тестовые наборы с учетом стратификации. Это означает, что при разделении сохраняется пропорциональное соотношение классов в целевой переменной. Это важно для данных финансового сектора, так как для них характерен большой дисбаланс классов.

<sup>8</sup>Метод кросс-валидации kFold разбивает исходный набор данных на k равных подмножеств (фолдов). В каждом из k итераций одно из подмножеств используется в качестве тестового набора, а остальные (k-1) – для обучения модели. После каждой итерации вычисляется метрика качества.

<sup>9</sup>TargetEncoder – метод кодирования категориальных переменных, при котором признаки заменяются смесью апостериорной вероятности цели с учетом конкретного категориального значения и априорной вероятности цели по всем обучающим данным.

ных методов Catboost обусловлено необходимостью преобразования данных для реализации методов снижения дисбаланса классов. На базовой модели методы снижения дисбаланса классов не применялись.

#### МЕТОДЫ СНИЖЕНИЯ ДИСБАЛАНСА КЛАССОВ

После разделения данных на каждом наборе данных, сплите, фолде на тренировочные, валидационные и тестовые к тренировочным данным применяется метод снижения дисбаланса. При этом к валидационным и тестовым данным метод снижения дисбаланса не применяется, потому что они должны отображать распределение реальных данных.

Для применения методов снижения дисбаланса применяется библиотека Python imblearn.

##### **Методы, увеличивающие количество наблюдений положительного класса:**

- Случайная избыточная выборка (RandomOverSampler). Генерация дополнительных наблюдений путем дублирования примеров из миноритарного класса (класса с меньшим количеством примеров) до выравнивания количества наблюдений миноритарного и мажоритарного классов.

- Synthetic Minority Oversampling Technique (SMOTE) [1]. Метод находит ближайших соседей миноритарного класса и генерирует образцы, делая линейную интерполяцию между ними.

- Adaptive Synthetic (ADASYN) [2] генерирует больше наблюдений в сложных областях, где алгоритм классификации испытывает трудности в разделении классов. Это позволяет модели лучше учиться на сложных частях пространства признаков.

- BorderlineSMOTE [3]. Этот алгоритм является вариантом оригинального алгоритма SMOTE. Для генерации наблюдений BorderlineSMOTE выбирает наблюдения на границе между двумя классами. Генерируется фиксированное число наблюдений для каждой точки миноритарного класса.

##### **Методы, снижающие количество наблюдений отрицательного класса:**

- Случайная недобырка (RandomUnderSampler). Наблюдения отрицательного класса выбираются случайно, по количеству наблюдений положительного класса.

- TomekLinks [4]. Суть метода состоит в том, что ищутся ближайшие соседи из разных классов. Если наблюдение из мажоритарного класса имеет ближайшего соседа из миноритарного класса, то наблюдение из мажоритарного класса считается неправильным и удаляется. За счет этого метода снижаются перекрытия классов и улучшаются показатели качества классификации.

- Правила очистки соседних наблюдений (NeighbourhoodCleaningRule) [5]. Метод направлен на улучшение качества данных. Для каждого наблюдения миноритарного класса определяются его ближайшие соседи. Если среди ближайших соседей много наблюдений отрицательного класса, то миноритарное наблюдение удаляется. Если экземпляр мажоритарного класса имеет много соседей из миноритарного класса, его также удаляют.

##### **Смешанные методы снижения дисбаланса классов:**

- Synthetic Minority Oversampling Technique + Edited Nearest Neighbors (SMOTEENN) [6]. Объединяет метод синтетического увеличения количества наблюдений миноритарного класса (SMOTE) и метод очистки данных (ENN), удаляющий наблюдения обоих классов, которые не поддерживаются ближайшими соседями.

- SMOTETomek [7]. На первом этапе генерируются синтетические наблюдения по методу SMOTE, на втором этапе данные очищаются по методу TomekLinks.

## ОБСУЖДЕНИЕ

В соответствии с описанной выше схемой валидации несколько раз проводилось разделение данных на тренировочный, валидационный и тестовый наборы данных. На тренировочных данных модель обучалась, валидационные данные использовались для настройки параметров модели машинного обучения и остановки процесса обучения, а на тестовых данных замерялись показатели качества работы алгоритма.

Первым этапом проводились обучение и настройка базовой модели машинного обучения без применения методов корректировки дисбаланса классов. Замеры показателей качества классификации проводились на тестовых данных, которые не использовались ни при обучении, ни при настройке алгоритма. Так как схема валидации предусматривает многократное разделение данных, результаты расчетов показателей качества классификации усреднялись.

Далее для каждого метода корректировки дисбаланса после применения разделения данных, согласно схеме валидации, на тренировочные, валидационные и тестовые к данным для тренировки применялся метод корректировки дисбаланса, и на этих данных обучалась базовая модель машинного обучения. К валидационным данным, используемым при настройке гиперпараметров модели машинного обучения и остановки процесса обучения, методы корректировки дисбаланса не применялись, чтобы свойства распределения признаков наборов данных, используемых при настройке параметров модели машинного обучения, по возможности, максимально совпадали с распределением для новых данных. После обучения модели проводились замеры показателей качества классификации на тестовых данных, которые не использовались ни при обучении, ни при настройке алгоритма. Так же, как и в базовой модели машинного обучения, результаты расчетов показателей качества классификации усреднялись.

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Таблица 2. Результаты экспериментов

Table 2. Results of the experiments

df №	Наименование метода	roc auc	balanced accuracy	avg precision	brier score	recall (sens-ty)	precision	f1 scr	specificity
1	Base	0,775	0,512	0,277	0,066	0,026	<b>0,674</b>	0,050	<b>0,999</b>
1	ADASYN	0,772	0,513	0,268	0,067	0,029	0,612	0,055	0,998
1	BorderlineSMOTE	0,773	0,514	0,268	0,067	0,030	0,606	0,057	0,998
1	NeighbourhoodCleaningRule	0,768	0,528	0,259	0,068	0,061	0,497	0,108	0,995
1	RandomOverSampler	<b>0,818</b>	<b>0,714</b>	<b>0,548</b>	<b>0,059</b>	0,459	0,572	<b>0,509</b>	0,970
1	RandomUnderSampler	0,765	0,696	0,243	0,200	<b>0,700</b>	0,166	0,269	0,692
1	SMOTE	0,772	0,514	0,268	0,067	0,029	0,609	0,055	0,998
1	SMOTEENN	0,768	0,552	0,256	0,071	0,119	0,421	0,185	0,986
1	SMOTETomek	0,773	0,514	0,269	0,067	0,030	0,608	0,057	0,998
1	TomekLinks	0,773	0,514	0,273	0,067	0,030	0,636	0,058	0,998
<b>1 Итого</b>		<b>0,776</b>	<b>0,557</b>	<b>0,293</b>	<b>0,080</b>	<b>0,151</b>	<b>0,540</b>	<b>0,140</b>	<b>0,963</b>

2	Base	0,764	0,533	0,344	0,090	0,071	<b>0,648</b>	0,128	<b>0,995</b>
2	ADASYN	0,763	0,534	0,344	0,090	0,073	0,640	0,131	0,995
2	BorderlineSMOTE	0,763	0,535	0,344	0,090	0,075	0,638	0,135	0,994
2	NeighbourhoodCleaningRule	0,757	0,563	0,325	0,094	0,147	0,480	0,225	0,979
2	RandomOverSampler	<b>0,809</b>	<b>0,725</b>	<b>0,553</b>	<b>0,086</b>	0,508	0,537	<b>0,522</b>	0,942
2	RandomUnderSampler	0,761	0,692	0,331	0,200	<b>0,692</b>	0,228	0,342	0,691
2	SMOTE	0,763	0,534	0,343	0,090	0,073	0,641	0,131	0,995
2	SMOTEENN	0,752	0,577	0,316	0,097	0,186	0,428	0,260	0,967
2	SMOTETomek	0,763	0,539	0,343	0,090	0,085	0,616	0,150	0,993
2	TomekLinks	0,763	0,540	0,340	0,090	0,087	0,608	0,152	0,993
<b>2 Итого</b>		<b>0,766</b>	<b>0,577</b>	<b>0,358</b>	<b>0,102</b>	<b>0,200</b>	<b>0,546</b>	<b>0,218</b>	<b>0,954</b>
3	Base	<b>0,901</b>	0,515	0,216	0,010	0,029	<b>0,820</b>	0,056	<b>1,000</b>
3	ADASYN	0,884	0,529	0,181	0,010	0,058	0,558	0,105	0,999
3	BorderlineSMOTE	0,886	0,532	0,183	0,010	0,065	0,533	0,116	0,999
3	NeighbourhoodCleaningRule	0,901	0,517	0,214	0,010	0,035	0,774	0,067	1,000
3	RandomOverSampler	0,846	0,673	<b>0,413</b>	<b>0,009</b>	0,348	0,654	<b>0,454</b>	0,998
3	RandomUnderSampler	0,899	<b>0,820</b>	0,163	0,127	<b>0,817</b>	0,049	0,092	0,823
3	SMOTE	0,884	0,529	0,181	0,010	0,058	0,559	0,106	0,999
3	SMOTEENN	0,885	0,539	0,179	0,010	0,078	0,486	0,135	0,999
3	SMOTETomek	0,884	0,529	0,181	0,010	0,059	0,563	0,107	0,999
3	TomekLinks	0,901	0,515	0,214	0,010	0,030	0,811	0,058	1,000
<b>3 Итого</b>		<b>0,887</b>	<b>0,570</b>	<b>0,212</b>	<b>0,021</b>	<b>0,158</b>	<b>0,581</b>	<b>0,130</b>	<b>0,982</b>

### Выводы

В результате проведенного исследования был получен рост показателей качества по сравнению с базовой моделью в случае применения метода RandomUnderSampler. Рост метрик ROC\_AUC и balanced accuracy свидетельствует о том, что с учетом несбалансированности классов повысилась ранжирующая способность модели и ее точность. Улучшение brier\_score указывает на то, что модель с применением RandomUnderSampler лучше откалибрована по сравнению с базовой моделью. Происходит также значительное увеличение метрики полноты (recall, sensitivity) и небольшое снижение метрики precision. Это говорит о том, что в целом качество модели улучшилось, но она стала генерировать больше ложноположительных предсказаний, что может быть критично при настройке функции принятия решения. В то же время высокий уровень специфичности specificity говорит о том, что базовая модель увереннее классифицирует отрицательные примеры по сравнению с моделью с применением метода RandomUnderSampler.

Значительный рост параметра полноты важен в тех областях, где критично выявить большинство положительных примеров (например, в медицине) несмотря на цену. В случае финансовых данных зачастую принятие решения об отказе в выдаче кредита происхо-

дит исходя из маржинальной прибыли групп наблюдений с определенным уровнем прогноза модели ИИ. Поэтому без учета роста маржинальной прибыли увеличение метрики полноты (sensitivity, recall) не является определяющим.

Лучшим показателем качества для финансовых несбалансированных данных, отражающим ранжирующую способность алгоритма, является метрика ROC\_AUC. Дело в том, что показатели precision и recall (f-меры) являются производными от ROC\_AUC и рассчитываются после назначения точки отнесения к положительному классу.

Интересен тот факт, что базовая модель и модель, построенная с применением метода RandomUnderSampler, значительно отличаются по структуре показателей качества. Применение методов объединения их прогнозов (стекинг) может привести к объединению их характеристик и еще большему росту показателей качества.

### ЗАКЛЮЧЕНИЕ

Рассмотрены методы учета дисбаланса классов в финансовых данных при построении моделей машинного обучения. Применение 9 методов учета дисбаланса к каждому из трех наборов данных по розничному кредитованию позволило провести сравнительный анализ метрик качества классификации для всех методов и определить методы, которые показали улучшение качества классификатора по сравнению с базовой моделью. Расчеты продемонстрировали улучшение показателей качества для метода RandomOverSampler. Результаты свидетельствуют о перспективности дальнейших исследований методов учета дисбаланса классов при исследовании финансовых данных, а также о целесообразности применения рассмотренных методов на практике.

### СПИСОК ЛИТЕРАТУРЫ / REFERENCES

1. Chawla N.V., Bowyer K.W., Hall L.O., Kegelmeyer W.P. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*. 2002. Vol. 16. Pp. 321–357. DOI: 10.1613/jair.953
2. He H., Bai Y., Garcia E.A., Li S. Adasyn: adaptive synthetic sampling approach for imbalanced learning. *In 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*. 2008. Pp. 1322–1328. DOI: 10.1109/IJCNN.2008.4633969
3. Han H., Wang W.-Y., Mao B.-H. Borderline-smote: a new over-sampling method in imbalanced data sets learning. *International conference on intelligent computing*. 2005. Pp. 878–887. Springer. DOI: 10.1007/11538059\_91
4. Tomek I. Two modifications of cnn. *IEEE Trans. Systems, Man and Cybernetics*. 1976. Vol. 6. Pp. 769–772. DOI: 10.1109/TSMC.1976.4309452
5. Laurikkala J. Improving identification of difficult small classes by balancing class distribution. *In Conference on Artificial Intelligence in Medicine in Europe*. 2001. Pp. 63–66. Springer. DOI: 10.1007/3-540-48229-6\_9
6. Batista G., Prati R.C., Monard M.C. A study of the behavior of several methods for balancing machine learning training data. *ACM Sigkdd Explorations Newsletter* 2004. Vol. 6. No. 1. Pp. 20–29. DOI: 10.1145/1007730.1007735
7. Batista G., Bazzan B., Monard M., Balancing Training Data for Automated Annotation of Keywords: a Case Study. *In WOB*. 2003. Pp. 10–18. BibTeX key: conf/wob/BatistaBM03

**Вклад авторов:** все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

**Contribution of the authors:** the authors contributed equally to this article. The authors declare no conflicts of interests.

**Финансирование.** Исследование проведено без спонсорской поддержки.

**Funding.** The study was performed without external funding.

### **Информация об авторах**

**Константинов Алексей Федорович**, аспирант кафедры информатики, Российский экономический университет им. Г. В. Плеханова;

115054, Россия, Москва, Стремянный переулок, 36;

konstantinovaf@gmail.com, ORCID: <https://orcid.org/0009-0000-9591-3301>, SPIN-код: 3088-3121

**Дьяконова Людмила Павловна**, канд. физ.-мат. наук, доцент, кафедра информатики, Российский экономический университет им. Г. В. Плеханова;

115054, Россия, Москва, Стремянный переулок, 36;

Dyakonova.LP@rea.ru, ORCID: <https://orcid.org/0000-0001-5229-8070>, SPIN-код: 2513-8831

### **Information about the authors**

**Alexey F. Konstantinov**, Post-graduate Student, Department of Informatics, Plekhanov Russian University of Economics;

115054, Russia, Moscow, 36 Stremyannyy lane;

konstantinovaf@gmail.com, ORCID: <https://orcid.org/0009-0000-9591-3301>, SPIN-code: 3088-3121

**Lyudmila P. Dyakonova**, Candidate of Physical and Mathematical Sciences, Associate Professor, Department of Informatics, Plekhanov Russian University of Economics;

115054, Russia, Moscow, 36 Stremyannyy lane;

Dyakonova.LP@rea.ru, ORCID: <https://orcid.org/0000-0001-5229-8070>, SPIN-code: 2513-8831