

УДК 336.66

В. В. Соловьев

Санкт-Петербургский государственный экономический университет,
Санкт-Петербург

Е. А. Синцова

Санкт-Петербургский государственный экономический университет, Санкт-Петербург;
Санкт-Петербургский университет технологий управления и экономики,
Санкт-Петербург, e-mail: sintsova_elena@hotmail.com

ПРИМЕНЕНИЕ МОДЕЛИ «СЛУЧАЙНЫЙ ЛЕС» ПРИ АНАЛИЗЕ ПРОИЗВОДСТВЕННЫХ БИЗНЕС-ПРОЦЕССОВ

Ключевые слова: анализ данных, бизнес-процессы, большие данные, машинное обучение, случайный лес, эффективность.

Актуальность статьи обусловлена увеличением объема данных, генерируемых промышленными предприятиями в реальном времени, и необходимостью выявления закономерностей, влияющих на эффективность производственных процессов. Традиционные методы анализа данных зачастую недостаточны для учета сложной природы этих данных, что создает потребность в современных технологиях, таких как машинное обучение. Цель исследования – исследование преимуществ и недостатков при формировании модели «Случайный лес» при использовании её на предприятиях, изучение её роли в повышении эффективности и оптимизации производственных систем. Основными задачами данного исследования являются: определить метод построения и методики работы модели «Случайный лес», выявить способы применения модели «Случайный лес» в условиях анализа производственных бизнес-процессов, выделить преимущества, недостатки и дать рекомендации применения модели «Случайный лес» в промышленности. Исследование основано на анализе теоретических и практических подходов к использованию модели «Случайный лес». Проведенное исследование показало высокую эффективность модели «Случайный лес» в прогнозировании и анализе производственных данных. Были выявлены ключевые преимущества, такие как устойчивость к переобучению, возможность работы с большими объемами данных и высокая точность предсказаний. Также отмечены недостатки, связанные с длительностью обучения и чувствительностью к отсутствию данных. Применение модели «Случайный лес» способно оказать положительное влияние на оптимизацию производственных процессов и принятие управленческих решений. Рекомендовано использовать качественные наборы данных и параллельное применение других моделей машинного обучения для повышения точности прогнозов.

V. V. Solovyev

St. Petersburg State University of Economics, St. Petersburg

E. A. Sintsova

St. Petersburg State University of Economics, St. Petersburg;
St. Petersburg University of Management Technologies and Economics,
St. Petersburg, e-mail: sintsova_elena@hotmail.com

APPLICATION OF THE «RANDOM FOREST» MODEL IN THE ANALYSIS OF PRODUCTION BUSINESS PROCESSES

Keywords: data analysis, business processes, big data, machine learning, random forest, efficiency.

The relevance of the article is due to the increase in the volume of data generated by industrial enterprises in real time and the need to identify patterns affecting the efficiency of production processes. Traditional data analysis methods are often insufficient to account for the complex nature of this data, which creates the need for modern technologies such as machine learning. The purpose of the study is to study the advantages and disadvantages of forming the “Random Forest” model when using it in enterprises, the study of its role in improving the efficiency and optimization of production systems. The main objectives of this study are to determine the method of construction and methods of operation of the “Random Forest” model, to identify ways of using the “Random Forest” model in the context of analysis of production business processes, to identify advantages, disadvantages and to make recommendations for the use of the “Random Forest” model in industry. Methodology: The research is based on the analysis of theoretical and practical approaches to the use of the “Random Forest” model. Results: The research

conducted has shown the high efficiency of the “Random Forest” model in forecasting and analyzing production data. Key advantages have been identified, such as resistance to overfitting, the ability to work with large amounts of data, and high prediction accuracy. Disadvantages related to the duration of training and sensitivity to missing data are also noted. Conclusions: The use of the “Random Forest” model can have a positive impact on the optimization of production processes and management decision-making. It is recommended to use high-quality datasets and parallel application of other machine learning models to improve the accuracy of forecasts.

Введение

На сегодняшний день промышленные предприятия генерируют огромные объемы данных в реальном времени по ходу своей производственной деятельности, в которых можно выявить определенные закономерности. Однако, существует большое множество возможных моделей, используемых в ходе анализа деятельности. Одной из таких является модель «Случайный лес», способная генерировать прогнозные результаты с целью принятия стратегически важных решений корпораций.

Прежде всего, выдвигается гипотеза о том, что применение модели «Случайный лес» для анализа производственных бизнес-процессов позволит выявить скрытые закономерности и зависимости между ключевыми показателями, что, в свою очередь, приведет к повышению эффективности производства и оптимизации бизнес-процессов.

За последнее десятилетие метод случайного леса стал популярным инструментом в области машинного обучения благодаря своей универсальности, точности и устойчивости к переобучению. Особенное распространение случайный лес нашел в задачах прогнозирования и анализа информации, играя значительную роль в принятии решений на основе обработанных данных.

Материалы и методы исследования

Литературные источники были выбраны на основании темы исследования, включая тематики цифровизации бизнеса, анализа данных и моделей машинного обучения.

При проведении исследования, в качестве исходных данных были использованы материалы научных статей по темам цифровизации бизнес-процессов в промышленности, анализа данных и моделей машинного обучения [1-4].

Важным аспектом подхода к анализу является обращение внимания на факторы, способные повлиять напрямую или косвенно на ожидаемый результат. Успешное применение данного подхода заклю-

чается в командной работе на всех её этапах согласования.

Наиболее популярными подходами в прогнозировании стали методы случайного леса и линейной регрессии. Среди публикаций выделяется статья Барбасовой Т.А., Бауман Е.В., Самолетовой П.А., Черепановой С.А. Авторы сравнивают три метода машинного обучения – линейную регрессию, случайный лес и градиентный бустинг для прогнозирования производительности на предприятии. Исследование показало преимущество метода случайный лес в точности прогнозов, подчеркнув их значимость для технико-экономического управления [5].

Работа Жукова П.И., Глущенко А.И., Фомина А.В. направлено на исследование методов машинного обучения. Было установлено, что алгоритм случайного леса показал заметно меньшую погрешность по сравнению с традиционной конечно-разностной моделью. Подобные исследования демонстрируют высокую эффективность применения машинного обучения, позволяя снижать энергозатраты и повышать качество продукции [6].

Отдельное внимание заслуживает статья Кулабуховой М.В посвященная моделированию состояния экономической безопасности предприятий с помощью случайного леса. Применяя ансамбль деревьев, автор получил точные прогнозы, снизив отклонение предсказанных значений ключевых индикаторов экономической безопасности. Такой подход позволит принимать экономически обоснованные решения и предотвращать потенциальные угрозы [7].

Наконец Шинкевич А.И., Малышева Т.В. предложили использовать ансамблевый метод случайного леса для диагностики зависимости основного производства от производственной инфраструктуры. Выделены классы производств по уровню зависимости, что открывает широкие возможности для последующей реструктуризации и оптимизации бизнес-процессов [8].

Работа авторов Ария М., Куккурулло К., Гнассо А. раскрывает комплексный анализ

техник расшифровки алгоритмов случайного леса. Методологическая база исследования сфокусирована на рассмотрении двух ключевых механизмов работы с древовидными структурами и агрегации узловых элементов, позволяющих извлекать закономерности из ансамблей решающих деревьев. Экспериментальная проверка выявила существенные расхождения в эффективности применения указанных подходов при обработке разнородных массивов информации, что подчеркивает специфические преимущества каждого метода. Полученные результаты значительно расширяют представления о возможностях интерпретации моделей случайного леса и перспективах их практического использования в разных сферах [9].

Современная модификация алгоритма случайного леса, разработанная группой исследователей под руководством Чен Дж., опирается на множественную мультиномиальную модель, обеспечивающую высокую согласованность результатов при сохранении точности вычислений. Практическое применение усовершенствованного метода позволяет достичь улучшенных прогностических показателей при решении задач классификации и регрессии, сохраняя базовую простоту классического алгоритма [10].

В статье авторов Ху Дж. и Шимчак С. представлен всесторонний анализ модификаций метода случайного леса применительно к обработке продольных данных. Практические исследования демонстрируют значительную результативность алгоритмических решений случайного леса при работе с повторяющимися измерениями, что открывает широкие возможности для применения в социальных и биомедицинских исследованиях [11].

В научной работе Лехнер М. и Окаса Г. внедрена модернизированная модификация алгоритма случайного леса применительно к упорядоченным моделям выбора. Результаты количественного анализа демонстрируют значительное превосходство методологии случайного леса над традиционными эконометрическими подходами в аспектах точности и стабильности получаемых результатов [12].

Результаты исследований Рамеш Р. и Джейкартик М. подтверждают эффективность применения метода случайного леса при построении прогнозных моделей спроса на продукцию. Модифицированный алгоритм случайного леса демонстрирует вы-

сокую точность идентификации значимых переменных и отсеивания второстепенных факторов, что существенно повышает качество управленческих решений в производственном планировании [13].

Статья Салман Х.А., Калакеч А., Стейти А. предлагает всесторонний обзор алгоритма случайного леса, объясняющий его концептуальные основы и методы использования. Авторы отмечают важность случайного леса в задачах классификации и прогнозирования, подчёркивая его гибкость и способность справляться с разнородными данными [14].

Алгоритм случайного леса продемонстрировал свою ценность в прогнозировании и принятии решений в различных областях: промышленность, материаловедение, биология, экономика и социальная наука. Современное состояние исследований подтверждает эффективность случайного леса как инструмента для прогнозирования, анализа данных и моделирования процессов. Дальнейшие исследования могут быть сосредоточены на повышении точности, надежности и интерпретации моделей случайного леса, а также на их адаптации к специальным случаям и нестандартным ситуациям.

Основными задачами данного исследования являются:

1. Определить метод построения и методики работы модели «Случайный лес»
2. Выявить способы применения модели «Случайный лес» в условиях анализа производственных бизнес-процессов
3. Выделить преимущества, недостатки и дать рекомендации применения модели «Случайный лес» в промышленности.

В задачах исследования основной акцент сделан на прогнозировании, диагностике и интерпретации сложных моделей, которые помогают принимать эффективные управленческие решения.

Наибольшее распространение модель случайного леса получила в экономике, металлургии, машиностроении, инженерии и энергетике, что связано с высокими требованиями к точности прогнозов и масштабом обрабатываемых данных.

Широкий спектр типов данных включает временные ряды, тексты, разные типы данных, что свидетельствует о гибкости и универсальности модели.

Цель исследования исследование преимуществ и недостатков при формировании

модели «Случайный лес» при использовании её на предприятиях, изучение её роли в повышении эффективности и оптимизации производственных систем.

Материал и методы исследования

Исследование основано на большом объеме информации, полученной из литературных источников и специализированных изданий, касающихся тем цифровизации бизнес-процессов, анализа данных и методов машинного обучения. Были выбраны статьи, относящиеся к исследованию и применению модели «Случайный лес» в промышленности, а также публикации, освещающие тему анализа производственных бизнес-процессов и финансовое управление предприятием.

Исследование построено на анализе современных подходов к анализу данных и машинному обучению. Изучение теоретических основ модели «Случайный лес», её истории и принципов работы, что позволило глубоко понять сущность и преимущества модели.

Детальное изложение особенностей алгоритма «Случайный лес», включающего случайный отбор признаков и данных при обучении, что уменьшает вероятность переобучения и повышает точность прогнозов.

Результаты исследования и их обсуждение

«Случайный лес» – метод машинного обучения, состоящий в формировании модели, позволяющей на основании значений некоторого набора независимых переменных предсказать значение дискретной либо непрерывной зависимой переменной. Модель представляет собой метод обучения, разработанный Лео Брэманом в 2001 году, суть модели заключается в использовании ансамблевого подхода, который улучшает общую модель взаимосвязанных элементов [2]. Само множество моделей называют набор деревьев решений, каждое дерево из которых обучается на случайном подмножестве данных, что позволяет избежать возникновения проблем с переобучением, которое может возникать при использовании одиночных деревьев. Данный метод гарантирует устойчивость и качество полученных предсказанных результатов. Если произвести сравнение на простых вещах, то можно сказать, что деревья решений являются строительным материалом самой модели,

при этом делая разбивку данных на основе определенных критериев, что позволяет извлечь определенный результат из полученных расчетов.

Одной из интересных особенностей «Случайного леса» является случайный отбор признаков при каждой его итерации, т.е. для создания каждого дерева выбирается не все доступные признаки, а только их подмножество. Этот подход не только повышает разнообразие этих деревьев, но и позволяет избежать корреляций между деревьями, что способствует улучшению общей точности модели. Кроме того, метод позволяет эффективно преодолевать недостатки, связанные с одиночными деревьями решений, такие как высокая чувствительность к шуму и возможным разбросам в данных.

Построение подобной модели не требует профессиональных навыков программирования, достаточно будет базовых умений, азов, а также наличия инструкции построения, в которой будет подробно описаны этапы формирования кода.

Для того, чтобы модель «Случайный лес» была применима для аналитики эффективности предполагаемых прогнозных значений, необходимо выгрузить следующую информацию:

- 1) объем производства и производительность;
- 2) объем отгрузки;
- 3) коэффициенты эффективного использования оборудования;
- 4) себестоимость готовой продукции;
- 5) стоимость передела;
- 6) прочие производственные показатели с учетом отрасли исследования.

Всю перечисленную информацию при необходимости можно представить по дням, однако, рациональней будет использование месячных данных за одинаковый промежуток времени, например, за 3 года. В случае успешного оцифровывания данных модель может получить достаточно качественной.

Также, хотелось бы разобрать практическое применение данной модели. Представим, что производственные рабочие компании продолжительной период времени добавляли определенное количество материалов. Однако, конъюнктура рынка изменилась, появилась необходимость в производстве других изделий. Математическая модель «Случайный лес» позволяет спрогнозировать вероятные колебания пропорций сырьевых компонентов в производственном

процессе. Персонал предприятия анализирует полученные прогнозные значения и вносит необходимые корректировки в параметры технологического процесса. Экономическая служба производит расчет влияния предложенных изменений на финансовые показатели организации. Модель формирует оптимальный баланс материальных затрат с учетом производственной себестоимости и стоимости технологических операций.

Результаты проведенного исследования подлежат представлению руководству организации для принятия обоснованных управленческих решений касательно модификации производственных стандартов. Анализ полученных данных позволяет рассмотреть комплексные предложения по оптимизации технологических процессов, пересмотру действующих нормативов и внедрению современных методов ресурсного менеджмента.

Таким образом, методы машинного обучения представляют собой инструмент для оптимизации производственных процессов в промышленности.

Модель «Случайный лес» отражает высокую точность предсказаний классификации за счет объединения множества результатов деревьев [8], однако сам процесс интерпретации полученных результатов может вызвать затруднения без определенной технологической подготовки. Именно поэтому был выделен процесс согласования результатов с технологами производства. При этом, хотелось бы отметить, что модель часто используется не только в промышленных предприятиях, но и в медицине.

Также, модель достаточно устойчива к переобучению и выбросам, несмотря на достаточно долгое время самого обучения. Благодаря случайному отбору подмножеств данных и признаков «Случайный лес» снижает риск переобучения, что делает её особенно полезной в условиях реальных производственных данных, однако обучение модели может занимать много времени, что может повлиять на скорость принятия решений.

Сама модель, в случае анализа больших объемов данных, который чаще всего происходит на предприятиях промышленности, потребляет большие вычислительные ресурсы. Для исключения технических ошибок требуется наличие мощного компьютера, способного обработать большие массивы данных (Big Data). В случае наличия

такого мощного компьютера повышается способность обработки большего количества признаков, что является преимуществом модели.

Помимо вышеперечисленного, хотелось бы остановить внимание на чувствительности модели к отсутствующим данным, которые способны уменьшить эффективность модели. Это может происходить, если компания не ведет учет и не загружает данные по ключевым параметрам. В случае появления фактора, не учтенного в модели ранее, могут глобально измениться результаты. Именно поэтому, на ранних этапах требуется формирования списка зависимых переменных, определение которого можно осуществить также, как и интерпретацию с технологами производства.

Касательно рекомендаций по использованию данной модели предприятиями, то хотелось бы выделить потребность в качественном сборе информации, который способен обеспечить полноту рассматриваемых параметров на каждом этапе производственного процесса.

Также, существует потребность в подтверждении достоверности данных, именно поэтому рекомендуется опираться на совместное применение с другими моделями машинного обучения, основанного на концепции создания ансамбля моделей, который последовательно улучшает предсказания, отказываясь от ошибок предыдущих моделей. Необходимо учитывать оценку обобщающей способности модели машинного обучения на новых данных, и настройку параметров, т.е. процесс поиска параметров, которые задаются до начала обучения модели и не могут быть обучены непосредственно из имеющихся данных.

Применение модели «Случайный лес» в бизнес-процессах открывает новые возможности для анализа и улучшения производственной эффективности. Возможности рассматриваемой модели в обработке информации и анализе больших объемов данных позволяют компаниям вовремя реагировать на волатильность рынка, а также организовывать прогноз возможных изменений с высокой точностью для повышения конкурентоспособности.

Применение модели «Случайный лес» в анализе производственных бизнес-процессов демонстрирует огромный потенциал для повышения эффективности предприятий. Результаты настоящего исследования

подтверждают целесообразность использования данного метода в промышленности для прогнозирования и анализа данных. Преимущества модели «Случайный лес»:

1. Высокая точность прогнозов благодаря ансамблевому подходу, объединяющему множество деревьев решений.

2. Устойчивость к переобучению, так как каждая итерация модели случайно выбирает подмножество признаков и данных.

3. Способность эффективно работать с разнородными и зашумленными данными, снижая влияние аномалий и искажений.

4. Отсутствие необходимости в значительных навыках программирования для построения модели.

Следует отметить в работе недостаточное изученные области, которые включают в себя следующее:

- хотя случайный лес обладает высокой точностью, он часто воспринимается как черный ящик, поскольку механизмы работы отдельных деревьев и выбор признаков трудно объяснить простыми словами. Необходимы дополнительные исследования для разработки методов визуализации и объяснения работы модели, что повысило бы доверие к ее выводам среди пользователей и экспертов;

- большинство исследований фокусируются на крупных наборах данных, тогда как малые выборки остаются недостаточно изученными, требуются эксперименты, направленные на выявление поведения случайного леса при ограниченном количестве наблюдений;

- случайный лес хорошо работает с полными и чистыми наборами данных, но проблемы возникают при наличии пропусков или шума, недостаточно изучено поведение модели при недостатке или плохом качестве данных, что важно исследовать в прикладных задачах;

- когда речь идет о применении случайного леса в крупномасштабных системах, возникает проблема высоких требований к ресурсам памяти и процессора, нужны дальнейшие исследования по разработке эффективных способов сокращения потребления ресурсов без потери точности;

- в быстро развивающейся среде, такой как финансовый рынок или ритейл, модели нуждаются в постоянном обновлении, нужно изучить адаптацию случайного леса к изменениям в структуре данных и сценариях применения.

Рекомендации для будущих исследований и практической применимости заключается в необходимости улучшения интерпретируемости модели путем разработки инструментов визуализации и описания работы отдельного дерева решений, что облегчит внедрение модели в критичные процессы и повысит доверие к прогнозам, кроме этого, необходимо минимизировать влияние предельных значений данных и их пропусков, улучшив таким образом общую стабильность и точность прогнозов. В свою очередь это повлечет систематические тесты на выборках разного размера, для определения минимально допустимого размера данных для получения результата и определения приемлемой точности при небольших наборах данных. Необходимы инструменты и библиотеки с простым интерфейсом для эксплуатации модели специалистами, а также изучение модели случайного леса с потоковыми и облачными сервисами, предоставляющими доступ к большим объемам постоянно поступающих данных, повышая оперативность и точность принимаемых решений.

Интерпретация результатов модели может вызывать сложности, так как механизм работы случайного леса носит характер черного ящика, т. е. сложной модели, логика работы которой труднодоступна для понимания и объяснения. Модель воспринимается как некий загадочный механизм, куда поступают входные данные, а на выходе получают результаты, но детально непосто объяснить, как именно получены эти результаты.

Интерпретация результатов, полученных с помощью модели «Случайный лес», иногда оказывается проблематичной. Однако высокая точность и надежность прогнозов делают данную модель востребованной, несмотря на недостаток прозрачности.

Модель «Случайный лес» можно успешно применять для прогнозирования изменений в сырье и материалах, оценки экономической эффективности производства и оптимизации бизнес-процессов. Корректировка производственных процессов на основе аналитических данных модели способствует оптимизации расходов и росту прибыли компании. Максимальная эффективность модели достигается при наличии подробных сведений о технологических параметрах производства на этапе внедрения. Интеграция алгоритма «Случайный лес» с альтернативными методами машинного обучения

существенно повышает качество прогнозирования показателей.

Алгоритм машинного обучения «Случайный лес» служит мощным катализатором развития производственной аналитики, позволяя компаниям существенно укреплять рыночные позиции за счет глубокого анализа и совершенствования технологических процессов.

Методика построения и функционирования модели «Случайный лес» раскрыта через призму анализа производственных бизнес-процессов промышленного сектора.

Проведенное исследование позволило установить специфику применения модели, выявить ключевые достоинства и ограничения методологии, сформировать практические рекомендации по внедрению. Результаты анализа подтвердили эффективность использования «Случайного леса» при оптимизации производственных процессов в промышленности. Внедрение подобных аналитических инструментов способствует более глубокому анализу и оптимизации процессов, что может значительно повысить эффективность корпораций отрасли в целом.

Библиографический список

1. Астраханцева И.А., Герасимов А.С. Прогнозирование региональной инфляции на основе гибридной модели машинного обучения: градиентный бустинг и случайный лес // Научные труды Вольного экономического общества России. 2023. Т. 243, № 5. С. 200-226. DOI: 10.38197/2072-2060-2023-243-5-200-226. EDN: JATZYG.
2. Волошин Т.А., Зайцев К.С., Дунаев М.Е. Применение адаптивных ансамблей методов машинного обучения к задаче прогнозирования временных рядов // International Journal of Open Information Technologies. 2023. Т. 11, № 8. С. 57-63. EDN: YORFMA.
3. Черепица А.Д., Синцова Е.А. Оцифровка, цифровизация цифровая трансформация бизнес-процессов // Цифровая экономика и финансы: Материалы Международной научно-практической конференции, Санкт-Петербург, 17–18 марта 2022 года. СПб.: Центр научно-производственных технологий «Астерион», 2022. С. 293-297. EDN: VSZDWT.
4. Синцова Е.А., Зайцев И.А. Особенности моделирования бизнес-процессов в условиях цифровой экономики // Цифровая экономика и финансы : Материалы Международной научно-практической конференции, Санкт-Петербург, 17–18 марта 2022 года. СПб.: Центр научно-производственных технологий «Астерион», 2022. С. 269-274. EDN: CAZNEB.
5. Барбасова Т.А., Бауман Е.В., Самолетова П.А., Черепанова С.А. Применение методов машинного обучения для прогнозирования производительности доменной печи // Известия Тульского государственного университета. Технические науки. 2023. № 9. С. 48-51. DOI: 10.24412/2071-6168-2023-9-48-49. EDN: TWSIHW.
6. Жуков П.И., Глущенко А.И., Фомин А.В. Сравнение модели конечных разностей и машинного обучения для задачи прогнозирования температуры заготовки, нагреваемой в проходной печи // Управление большими системами: сборник трудов. 2022. № 95. С. 79-100. DOI: 10.25728/ubs.2022.95.5. EDN: SRADLH.
7. Кулабухова М.В. Применение алгоритма случайного леса в моделировании состояния экономической безопасности экономических систем // Образование. Наука. Научные кадры. 2023. № 4. С. 139-143. DOI: 10.24412/2073-3305-2023-4-139-143. EDN: CRBCXX.
8. Шинкевич А.И., Малышева Т.В. Диагностика зависимости основного производства от производственной инфраструктуры на основе ансамблевого метода машинного обучения // Общество: политика, экономика, право. 2023. № 8(121). С. 114-120. DOI: 10.24158/pep.2023.8.14. EDN: MEXXWQ.
9. Aria M., Cuccurullo C., Gnasso A. A comparison among interpretative proposals for Random Forests // Machine Learning with Applications. 2021. Vol. 6. P. 100094. DOI: 10.1016/j.mlwa.2021.100094. EDN: GLLFZQ.
10. Chen Ju., Wang X., Lei F. Data-driven multinomial random forest: a new random forest variant with strong consistency // Journal of Big Data. 2024. Vol. 11, No. 1. P. 34. DOI: 10.1186/s40537-023-00874-6. EDN: WWTTHO.
11. Hu J., Szymczak S. A review on longitudinal data analysis with random forest // Briefings in Bioinformatics. 2023. Vol. 24, No. 2. DOI: 10.1093/bib/bbad002. EDN: UUMLBC.
12. Lechner M., Okasa G. Random Forest estimation of the ordered choice model // Empirical Economics. 2024. DOI: 10.1007/s00181-024-02646-4. EDN: GNXUAV.
13. Ramesh Ranjitha, Jeyakarthic M. Predictive Analytics For Steel Market Using Improved Random Forest // Webology. 2021. Vol. 18. № 4824.
14. Salman H.A., Kalakech A., Steiti A. Random Forest Algorithm Overview // Babylonian Journal of Machine Learning. 2024. Vol. 2024. P. 69-79. DOI: 10.58496/bjml/2024/007. EDN: HWNARA.