

Программные системы и вычислительные методы*Правильная ссылка на статью:*

Никитин П.В., Андриянов Н.А., Горохова Р.И., Бахтина Е.Ю., Долгов В.И., Коровин Д.И. — Методология оценки рисков выполнения государственных контрактов с использованием средств машинного обучения // Программные системы и вычислительные методы. – 2023. – № 4. – С. 44 - 60. DOI: 10.7256/2454-0714.2023.4.44113 EDN: HMHCXC URL: https://nbppublish.com/library_read_article.php?id=44113

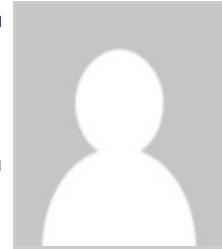
Методология оценки рисков выполнения государственных контрактов с использованием средств машинного обучения**Никитин Петр Владимирович**

ORCID: 0000-0001-8866-5610

кандидат педагогических наук

доцент, Департамент анализа данных и машинного обучения, Финансовый университет при Правительстве Российской Федерации

125993, Россия, г. Москва, Ленинградский проспект, 49

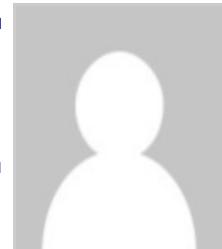
[✉ pvnikitin@fa.ru](mailto:pvnikitin@fa.ru)**Андрянов Никита Андреевич**

ORCID: 0000-0003-0735-7697

кандидат технических наук

доцент, Департамент анализа данных и машинного обучения, Финансовый университет при Правительстве Российской Федерации

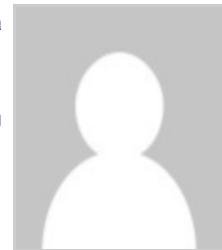
125163, Россия, г. Москва, ул. Ленинградский пр-т, 49

[✉ naandriyanov@fa.ru](mailto:naandriyanov@fa.ru)**Горохова Римма Ивановна**

кандидат педагогических наук

доцент, Департамент анализа данных и машинного обучения, Финансовый университет при Правительстве Российской Федерации

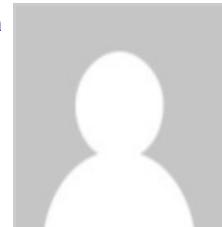
125167, Россия, г. Москва, Ленинградский проспект, 49

[✉ rigorokhova@fa.ru](mailto:rigorokhova@fa.ru)**Бахтина Елена Юрьевна**

кандидат физико-математических наук

Доцент, Московский автомобильно-дорожный государственный технический университет (МАДИ)

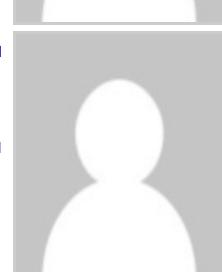
125319, Россия, г. Москва, Ленинградский проспект, 46

[✉ elbakh@gmail.com](mailto:elbakh@gmail.com)**Долгов Виталий Игоревич**

кандидат физико-математических наук

доцент, Департамент анализа данных и машинного обучения, Финансовый университет при Правительстве Российской Федерации

125319, Россия, г. Москва, ул. Ленинградский Проспект, 49

[✉ vidolgov@fa.ru](mailto:vidolgov@fa.ru)

Коровин Дмитрий Игоревич

доктор экономических наук

профессор, Департамент анализа данных и машинного обучения, Финансовый университет при Правительстве Российской Федерации

125319, Россия, г. Москва, ул. Ленинградский Проспект, 49

✉ dikorovin@fa.ru



[Статья из рубрики "Базы знаний, интеллектуальные системы, экспертные системы, системы поддержки принятия решений"](#)

DOI:

10.7256/2454-0714.2023.4.44113

EDN:

HMHCXC

Дата направления статьи в редакцию:

24-09-2023

Дата публикации:

22-11-2023

Аннотация: Предметом исследования является разработка программного комплекса для интеллектуального прогнозирования выполнения государственных контрактов с использованием методов машинного обучения и анализа неструктурированной информации. Объектом исследования является процесс контроля и принятия решений в области государственных закупок, включая выбор подрядчиков, выполнение контрактов и оценку сроков и стоимости их реализации. Особое внимание в исследовании уделяется разработке и применению интерпретируемых методов машинного обучения для решения задач оценки рисков выбора недобросовестного подрядчика, рисков невыполнения контракта в срок и прогнозирования вероятных сроков и стоимости реализации контракта. Авторы подробно рассматривают такие аспекты, как уникальный набор данных, который был собран из различных информационных систем. Они также разработали автоматизированные системы сбора и обновления данных, которые могут быть установлены на серверах заказчиков. В работе применялись методы машинного обучения, анализа неструктурированной информации и интерпретируемых методов. Были построены интерпретируемые модели машинного обучения для оценки риска выбора недобросовестного подрядчика, оценки риска невыполнения контракта в срок, а также оценки вероятных сроков и стоимости реализации контракта. В работе был собран уникальный набор данных, включающий более 83 тысяч данных по более чем 190 признакам с различных систем, таких как Реестр государственных закупок Единой информационной системы (ЕИС), Реестр недобросовестных поставщиков (РНП) ЕИС и Информационная система СПАРК. Разработаны автоматизированные системы сбора и

обновления данных, которые могут быть развернуты на серверах заказчиков. В ходе исследования были разработаны программные комплексы для интеллектуального прогнозирования выполнения государственных контрактов, которые предоставляют возможность проведения более точного анализа рисков с использованием методов анализа неструктурированной информации, моделей машинного обучения и интерпретируемых методов. Это позволяет повысить эффективность контроля выполнения государственных контрактов и снизить вероятность коррупции и нарушений. Исследование демонстрирует значимость и применимость методов и моделей машинного обучения в области государственных контрактов и предоставляет новые возможности для улучшения процессов контроля и принятия решений в сфере государственного закупочного дела.

Ключевые слова:

государственные контракты, машинное обучение, интеллектуальные системы, визуализация данных, метод опорных векторов, оценка рисков, прогнозирование, регрессионный анализ, интерпретируемый искусственный интеллект, ансамблевые методы

Статья подготовлена в рамках государственного задания Правительства Российской Федерации Финансовому университету на 2023 год

Введение

Вопросы контрактной деятельности могут иметь разную степень важности в зависимости от контекста и конкретной ситуации. Важными являются целый ряд аспектов, среди которых обязательными являются определение целей и обязательств сторон; условия оплаты; гарантии и ответственность сторон в контракте; конфиденциальность и защита данных; разрешение споров в случае возникновения разногласий между сторонами; условия изменения и расторжение контракта, являющееся важным для флексибильности и адаптации к изменяющимся обстоятельствам; соблюдение законодательства и регулирования. Важность этих вопросов может варьироваться в зависимости от типа контракта, отрасли и места осуществления контрактной деятельности. При подготовке и заключении контрактов, чтобы обеспечить ясность, защиту интересов сторон и снизить потенциальные риски рекомендуется тщательно рассмотреть все эти вопросы.

А. Я. Геллер заключает, что для снижения риска расторжения контрактов необходимо улучшить систему закупок, внедрить эффективные механизмы контроля и надзора, а также повысить квалификацию участников закупочных процедур. Исследование статьи предоставляет полезную информацию о причинах расторжения контрактов в системе государственных и муниципальных закупок Российской Федерации, что может быть полезно для оценки и управления рисками при осуществлении государственных контрактов. [\[1\]](#)

В работе М. М. Золотухиной и Н. А. Половниковой [\[2\]](#) выполнено исследование рисков, с которыми сталкиваются предприятия при выборе поставщиков и заключении контрактов. Отмечается, что неправильный выбор поставщиков и непродуманные условия контракта могут привести к серьезным финансовым и операционным проблемам, включая высокие цены, плохое качество товаров или услуг, несоблюдение сроков поставки и т.д. В статье приводятся различные виды рисков, такие как финансовые риски, операционные риски,

риски связанные с недостаточными знаниями о поставщиках и их репутацией, а также риски, связанные с изменениями в мировой и местной экономике. Для управления указанными рисками предлагается использовать анализ финансового состояния поставщика, мониторинг и оценка репутации поставщика, использование различных видов контрактов и т.д.

Статья Ю. Б. Гендлиной и др. посвящена исследованию рисков, связанных с заключением строительных контрактов с муниципальными заказчиками. Авторы уделяют внимание влиянию муниципального контекста на риски и особенности управления ими и предлагают рекомендации и стратегии по управлению рисками, связанными со строительными контрактами. Рассматриваются такие инструменты, как подробный анализ условий контракта, регулярный мониторинг и контроль выполнения работ, использование страхования и гарантий, а также разумное управление изменениями и рисковым контролем. [\[3\]](#)

В исследовании М. Ю. Алейниковой и Д. А. Голованова рассмотрены модели улучшения системы внутреннего контроля при осуществлении государственных закупок в Российской Федерации. В статье рассматриваются различные модели совершенствования системы внутреннего контроля, используемые в России, предложены решения, основанные на процессном подходе, риск-ориентированном подходе, использовании информационных технологий и автоматизации, а также на усилении контроля и независимости внутренних служб. В статье также представлен анализ преимуществ и недостатков каждой модели, а также опыт и результаты их применения в практике государственных закупок.

В статье Д. А. Елисеева и Д. А. Романова исследуются вопросы применения машинного обучения для прогнозирования рисков в госзакупках [\[5\]](#). Представлен анализ различных факторов (финансовая устойчивость поставщиков, репутация компаний, история выполнения контрактов и другие параметры) для выявления возможных рисков и проблем. Проведен сравнительный анализ машинного обучения и традиционных статистических методов для прогнозирования, который позволил выявить более высокую точность и эффективность прогнозирования рисков, которую можно достичь с помощью машинного обучения.

Исследование Ю. В. Немцева и О. Б. Миронец посвящено анализу развития рынка информационных технологий в области государственных закупок. Авторы рассматривают риски, возникающие во время взаимодействия участников рынка в конкурентной среде государственных закупок. Они предлагают систематическое классифицирование рисков и различные методы для оценки величины этих рисков. В своей работе они также изучают подходы и методы управления рисками, которые имеют важное значение для менеджмента. [\[6\]](#)

Применение теоретико-игровой модели Штакельберга для анализа затрат и рисков нарушения договора в российском минерально-сырьевом комплексе приведено в исследовании С. М. Лавлинского, А. А. Панина и А. В. Плясунова [\[7\]](#). В данной статье предпринимается попытка учесть институциональные особенности процесса формирования инвестиционного климата. Для этого в модели вводятся понятие трансакционных издержек и создается механизм страхования риска нарушения договора. Модель описывается как двухуровневая задача математического программирования, и для ее решения разработаны эффективные алгоритмы на базе метаэвристик.

Риски невыполнения государственных контрактов можно снизить на этапе предварительного отбора для участия в конкурсах перед заключением государственных контрактов. Для этой цели значимыми инструментами являются современные информационно-аналитические системы, реализуемые в сети Интернет. В. В. Скобелев [8] проводит анализ серьезной проблемы и рассматривает методы проверки участников конкурса проведения закупок. Методы машинного обучения и интеллектуальный анализ данных становятся важным инструментом и при работе с государственными закупками [9, 10]. Ю. М. Бекетнова [11] проводит анализ контрактов методами машинного обучения с целью выявления нарушений по нецелевым доходам, а А. З. Манцкава [12] в своем исследовании алгоритма машинного обучения использует для классификации завершенных государственных закупок.

Целый ряд исследований [13-16] посвящены исследованию гражданско-правовых рисков при заключении контрактов для нужд силовых ведомств, минимизации рисков и выработке системы мер по парированию рисков.

Изучение различных рисков в сфере закупок и заключении контрактов является важным аспектом в различных отраслях экономики. Особое внимание уделяется выбору поставщика, осуществляющему на конкурсной основе [1, 17]. Анализ финансовой устойчивости, репутации и качества услуг, товаров и надежности поставок поставщика являются ключевыми факторами при выборе. В различных статьях рассматриваются риски выполнения контрактов в разных сферах, таких как строительство, информационные технологии и минерально-сырьевой комплекс [3, 18, 6, 8]. Л. С. Травкина, П. В. Лисин, Е. А. Межевикина исследуют риски поставки продукции, не соответствующей требованиям контракта, что также приведет к его не выполнению [19]. Различные методы и подходы к управлению рисками, а также модели для повышения контроля и предотвращения нарушений контрактного законодательства все в большей мере реализуются при создании информационных систем для сферы государственных услуг в части разработки средств оценки проектных рисков [20].

Анализируя проведенные исследования, можно сделать вывод о текущей актуальности исследования рисков исполнения государственных контрактов. Существует большое количество работ, посвященных конкретным областям экономики и сырьевым отраслям, а также финансовым и гражданско-правовым рискам, аспектам связанных с поставщиками и их репутацией, а также влиянию изменений в мировой и местной экономике.

Из проведенного анализа можно выделить некоторые методы и средства, которые могут быть использованы для анализа и управления рисками при выполнении государственных контрактов. Среди них: методы управления рисками, детальный анализ условий контракта, регулярный мониторинг и контроль выполнения работ, а также использование страхования и гарантий. Однако, недостаточно исследований, которые системно рассматривают все риски и применяют автоматизированные интеллектуальные системы для их анализа.

В данной работе было проведено исследование, нацеленное на комплексный анализ и управление рисками при выполнении государственных контрактов. В рамках исследования предложены методы и подходы, основанные на применении машинного обучения и интеллектуального анализа данных, которые позволяют рассмотреть риски системно и эффективно. Данное исследование заполняет пробел в существующей литературе и предлагает новый подход к управлению рисками в данной области.

Методология разработки интеллектуального комплекса оценки рисков

В основу методологии автоматизированного интеллектуального контроля исполнения государственных контрактов заложены следующие положения:

1. Для успешного построения интеллектуальной автоматизированной системы контроля государственных контрактов необходим большой набор данных, содержащий максимально полезную информацию о предмете договора, финансовые и юридические сведения по компаниям-поставщикам и другую полезную информацию.
2. Интеллектуальная автоматизированная система контроля государственных контрактов должна быть направлена на решение трех задач: оценка риска выбора недобросовестного подрядчика; оценка риска невыполнения контракта в срок; оценка вероятных сроков и стоимости реализации контракта.
3. Интеллектуальная система, направленная на оценку риска выбора недобросовестного подрядчика, риска невыполнения контракта в срок и оценку вероятных сроков и стоимости реализации контракта, должна учитывать важность признаков при оценке и принятии решений. Важность признаков указывает на то, как сильно определенный признак или фактор влияет на риск или оцениваемую характеристику. Учет важности признаков позволяет системе сосредоточиться на наиболее информативных факторах и принимать во внимание их влияние на риски и оцениваемые характеристики. Это может помочь снизить шум и улучшить качество оценок и прогнозов, что в свою очередь улучшит принятие решений в контексте выбора подрядчиков, выполнения контрактов и оценки сроков и стоимости их реализации.
- 4 . После апробации результатов при оценке рисков исполнения государственных контрактов, интеллектуальная автоматизированная система контроля государственных контрактов может быть развернута на серверах потенциальных заказчиков.

Рассмотрим каждый из этапов.

Набор данных для оценки рисков исполнения государственных контрактов

Один из первых шагов в разработке методологии – сбор и структурирование данных, связанных с исполнением государственных контрактов. Данные будут включать в себя информацию о контрактах, подрядчиках, выполненных работах, сроках и оценке качества. Важно, чтобы данные были структурированы и удобно доступны для использования.

В рамках исследования собран уникальный набор данных, состоящий из более 83 тысяч данных по более чем 190 признаков с систем: Реестра государственных закупок Единой информационной системы (ЕИС) (<https://zakupki.gov.ru/>); Реестра недобросовестных поставщиков (РНП) ЕИС (<https://zakupki.gov.ru/epz/dishonestsupplier/search/results.html>); Информационной системы СПАРК (<https://spark-interfax.ru/>).

Более подробно набор данных для исследования описан в статье «Методика оценки важности признаков при анализе выполнения государственных контрактов», журнала «Национальная безопасность» [\[21\]](#).

Модуль оценки риска невыполнения контракта в срок

Данный модуль представляет собой инструмент для решения задачи классификации, обладающий функциональностью прогнозирования не только дискретных ответов

(«выполнится/не выполнится»), но и вероятностей возникновения соответствующих рисков. Такой подход обеспечивает гибкость системы. Модуль разработан на языке программирования Python в среде Jupyter Notebook. Скрипты сохранены в формате .ipynb и могут быть интегрированы в облачные сервисы Google Colab и Yandex Colab. Модуль поддерживает импорт, анализ, предобработку и классификацию данных, обучение и тестирование моделей, построение вероятностных распределений и визуализацию результатов для интерпретации результатов и конкурентного преимущества.

Проверка адекватности и качества работы модуля была выполнена путем анализа данных, собранных в трех информационных системах, охватывающих около 80 тысяч государственных контрактов по 192 признакам.

В процессе оценки рисков, переменная «Статус» была определена как целевая. Она может принимать два значения: «Контракт исполнен» и «Контракт прекращен», что сводит задачу к классификации. Исходный датасет содержит 83834 наблюдений и 192 переменных. Несмотря на высокую размерность и пропущенные значения в большинстве наблюдений, такой объем данных может способствовать качественной оценке вероятности расторжения контракта.

Распределение контрактов в датасете можно считать сбалансированным, что подтверждается рисунком 1, где «1» соответствует расторгнутым контрактам, а «0» – исполненным.

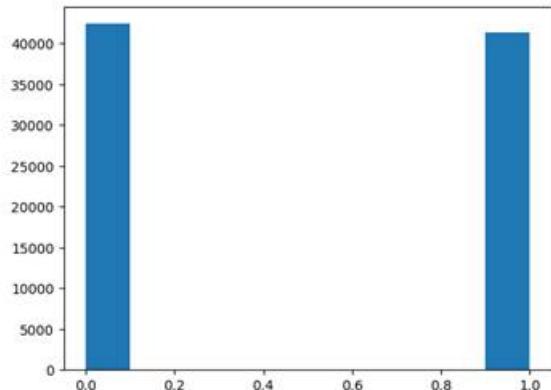


Рисунок 1 – Распределение целевого класса [Составлено авторами]

Для эффективной обработки данных будет применен метод меток кодирования для всех нецелочисленных признаков. В контексте большого объема данных, данный подход предпочтительнее one-hot кодирования. Кодирование будет проведено для каждого соответствующего столбца, а пропущенные значения будут заменены кодом 100000, обеспечивая получение нового кода даже при наихудших условиях кодирования.

Для прогнозирования вероятностей статуса «Контракт прекращен» будет обучена модель логистической регрессии. Обучающая и тестовая выборки были разделены в соотношении 85% к 15%.

После кодирования данных, значения в большинстве столбцов распределены в диапазоне от 0 до N, где N – количество категорий, в то время как числовые столбцы остаются неизменными. Для оптимизации эффективности модели, рекомендуется проведение стандартизации данных, что обеспечит равномерное распределение признаков.

На тестовых данных модель логистической регрессии показала метрику Accuracy, равную 97,01%. Это достаточно хороший результат, так как только 2,99% получают неверные метки при прогнозировании на новых данных.

Далее мы исследуем распределение вероятностей нарушения контракта в тестовой выборке с использованием оценки логистической регрессии (см. рисунок 2).

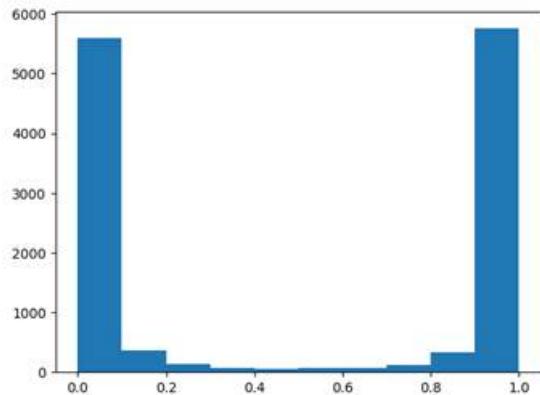


Рисунок 2 – Распределение вероятностей рисков контракта по логрегрессии [Составлено авторами]

На основе данных из рисунка 2, модель в большинстве случаев предоставляет определенные прогнозы, с вероятностями, приближающимися к 0 или 1. Процентное распределение в диапазоне от 0,1 до 0,9 является незначительным. Это указывает на ограниченность модели логистической регрессии в оценке вероятностей рисков. Аналогичный недостаток в гибкости ответов наблюдается в модели «Дерево решений», что подтверждается точностью модели, составляющей 97,89%, и иллюстрируется на рисунке 3.

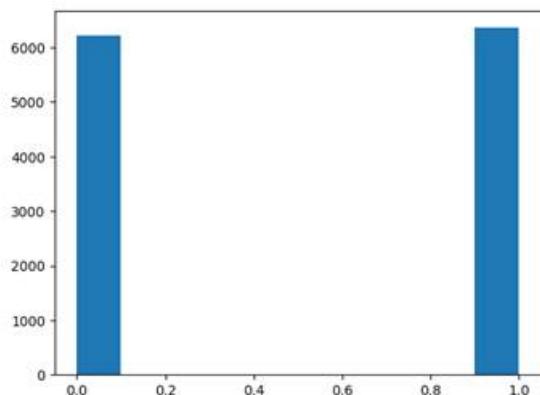


Рисунок 3 – Распределение вероятностей рисков контракта по дереву решений [Составлено авторами]

Наконец, построим модель на базе метода опорных векторов. Метрика точности на тестовых данных для этой модели составляет 97,88%, что незначительно уступает показателю модели «Дерево решений». Однако данный метод обеспечивает более гибкую оценку рисков. Распределение вероятностей рисков для данного метода представлено на рисунке 4.

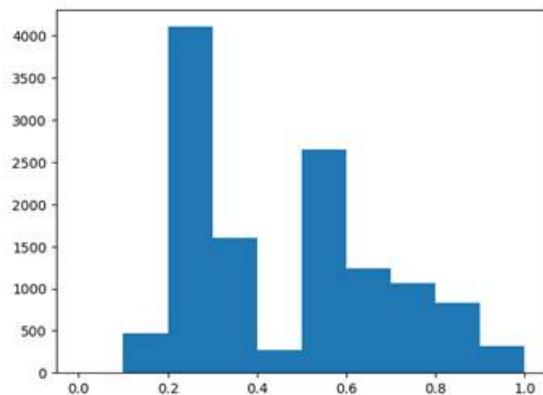


Рисунок 4 – Распределение вероятностей рисков контракта по методу опорных векторов
[Составлено авторами]

Анализ гистограммы показывает, что вероятности рисков по контрактам не обладают бимодальностью по границам и в целом распределены неравномерно. При этом большая часть контрактов классифицируется как «Контракт прекращен» с вероятностью, приближенной к 0,5. В таблице 1 представлены результаты сравнения различных моделей классификации.

Таблица 1 – Результаты сравнения разных моделей классификации

Модель	Точность	Вероятности рисков
Логистическая регрессия	97,01%	Либо имеют очень высокие значения, либо очень низкие
Дерево решений	97,89%	Только однозначные «да» и «нет»
Метод опорных векторов	97,88%	Распределены по всему интервалу от 0 до 1.

Если целью является детальный интеллектуальный анализ контрактов, наиболее подходящей моделью будет SVM, обеспечивающая прогнозирование различных вероятностей, соответствующих реальным условиям. Для SVM будет проведена оценка важности признаков. Наилучшие и наихудшие значения, определенные с помощью метода `feature importants`.

Из анализа мы можем увидеть, что важный вклад в прогноз вносят такие признаки, как налоговый орган компании, коэффициент покрытия процентов по EBIT, описание объекта закупки, коэффициент концентрации заемного капитала. Воздействие анализируемых факторов на выполнение госконтракта ассоциируется с разными аспектами: финансовой устойчивостью компании, определенной по налоговому статусу; коэффициентом покрытия процентов по EBIT; качеством описания закупаемого объекта; коэффициентом концентрации заемного капитала. Оценить влияние этих элементов на риск невыполнения госконтракта возможно через финансовый анализ, аудит, просмотр предыдущих контрактов, систему оценки рисков и моделирование.

Модуль оценки вероятных сроков и стоимости реализации контракта

Прогнозирование стоимости и сроков исполнения госконтрактов представляет собой задачу регрессии, для решения которой был создан модуль на Python, хранящийся в формате IPython Notebook (.ipynb) и совместимый с облачными платформами Google и Yandex. Данный модуль использует библиотеки Python: Pandas, Matplotlib и Scikit Learn для обработки, визуализации данных и обучения регрессионных моделей. Модуль обеспечивает возможность загрузки и обработки данных, разделение набора данных по

стоимости и времени выполнения, кодирование признаков, стандартизацию, тестиирование регрессионных моделей и визуализацию ключевых факторов, влияющих на сроки и стоимость контрактов.

Адекватность и точность разработанного модуля оценивались на основе данных из трех информационных систем, включающих около 80 тысяч контрактов (191 признак). Зависимыми переменными были определены стоимость и срок исполнения контракта, вычисленный как разница между датами начала и окончания. Были разработаны модели с ключевыми метриками R-квадрат, MSE и MAPE. Предобработка данных для классификации включала типизацию, LabelEncoding категориальных данных и стандартизацию для оптимизации алгоритмов. Исходный набор данных (83834 x 191) был уменьшен до одной переменной, представляющей разницу между двумя датами. Пропуски в данных были заполнены средними значениями по столбцу. Все данные были разделены в соотношении 85% на 15% для обучения и тестирования модели. Однако был замечен значительный эффект переобучения моделей. Результаты прогнозирования сроков исполнения контракта представлены в таблице 2.

Таблица 2 – Анализ моделей регрессии сроков исполнения контрактов

Модель	R-квадрат	MSE	RMSE	MAPE
Линейная регрессия	0.181	51077	226	0.436
Метод опорных векторов	0.153	52801	229	0.444
Метод опорных векторов (без стандартизации)	0.028	60621	246	0.537
Дерево решений	0.178	51250	226	0.556

Анализ представленной информации указывает на превосходство дерева решений в достижении оптимальных результатов, при условии стандартизации, без которой качество моделей ухудшается. Для дополнительного улучшения рекомендуется проведение более глубокого анализа признакового пространства и применение нейронных сетей. Эти выводы согласуются с результатами, полученными для моделей прогнозирования стоимости контракта, которые представлены в таблице 3.

Таблица 3 – Анализ моделей регрессии стоимости контрактов

Модель	R-квадрат	MSE	RMSE	MAPE
Линейная регрессия	-0.0004	430381086	20745,62812	0.506
Метод опорных векторов	0	430718896	20753,76824	0.502
Метод опорных векторов (без стандартизации)	-0.0004	430391492	20745,87892	0.527
Дерево решений	0.683	136382258	11678,28147	0.328

Анализ результатов указывает на превосходство дерева решений в прогнозировании стоимости контрактов, в то время как эффективность остальных моделей оставляет желать лучшего. Модель регрессии позволяет провести первичную оценку сроков и стоимости контракта, однако требует дополнительной корректировки. Отмечено переобучение моделей, которые показывают 100% результаты на тренировочной выборке, что предполагает необходимость применения методов регуляризации в будущем.

Отметим так же, что в данных находятся «выбросные» контракты, которые сильно долго

длятся по сравнению с другими. Это, вероятно, приводит к высоким ошибкам при прогнозировании.

Данные о сроках исполнения контрактов содержат выбросы и будут очищены согласно условию:

$$Q1 - 1,5R < Days < Q3 + 1,5R,$$

где $Q1$ и $Q3$ – первый и третий квартили ряда $Days$, R – межквартильный размах.

Данные по стоимости госконтрактов распределены равномерно (рис. 5).

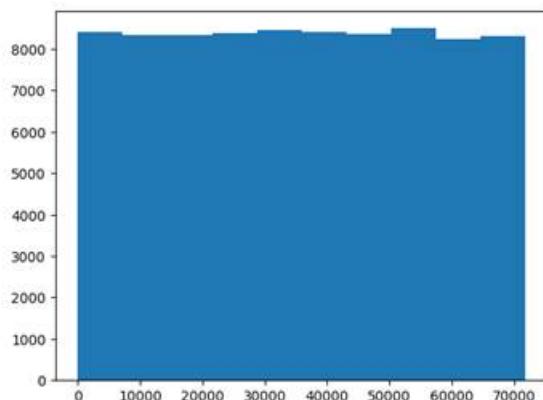


Рисунок 5 – Гистограмма стоимостей контрактов [Составлено авторами]

Исходя из эффективности деревьев решений, будут разработаны ансамблевые модели, включая Extra Tree и Random Forest. После удаления выбросов, обучающие и тестовые данные будут разделены в пропорции 90% к 10%. Результаты прогнозирования сроков контракта для различных ансамблей представлены в таблице 4.

Таблица 4 – Анализ моделей регрессии сроков исполнения контрактов

Ансамбль	R-квадрат	MSE	RMSE	MAPE
Gradient Boosting	0.372	13287	115	0.365
AdaBoost	0.195	17016	130	0.394
ExtraTree	0.960	840	28	0.027
Random Forest	0.888	2377	48	0.158

Из таблицы 4 видно, что и наиболее точные прогнозы дают модели ExtraTree и RandomForest. При этом средняя ошибка при использовании ансамбля ExtraTree не превышает одного месяца, что является удовлетворительным результатом.

В таблице 5 показаны результаты прогнозирования стоимости контракта.

Таблица 5 – Результаты прогнозирования стоимости контракта

Ансамбль	R-квадрат	MSE	RMSE	MAPE
Gradient Boosting	0.697	129699489	11388	0.270
AdaBoost	0.310	294692017	17166	0.414
ExtraTree	0.950	21289343	4614	0.034
Random Forest	0.952	20454948	4522	0.029

Модели ExtraTree и Random Forest демонстрируют оптимальные результаты в прогнозировании стоимости, со средней ошибкой в 4500-4600 рублей, что считается

приемлемым.

Из модели можно выявить, что идентификация заказчика, его требования и условия контракта, эффективность организаций в процессах закупки и контроля, номер извещения о торгах, описание объекта закупки и идентификационный код закупки влияют на сроки исполнения контракта. Эти факторы могут напрямую или косвенно влиять на прогноз сроков исполнения контракта. Для более точного прогноза необходимо учесть и другие факторы.

Проведем детальный анализ факторов, существенно влияющих на стоимость государственного контракта:

1. Объем и стоимость поставляемой продукции или услуг, отраженные в числе заполненных записей. Большее количество записей обычно свидетельствует о большем объеме поставок, что влияет на общую стоимость контракта. Данный фактор также может влиять на расчеты, связанные с доставкой, хранением и логистикой.
2. Код позиции объекта закупки, который определяет конкретные продукты или услуги для закупки. Различные объекты закупки могут иметь различные стоимости в зависимости от уровня технологичности или специфических навыков, необходимых для их производства или предоставления.
3. Общая стоимость поставленной продукции, которая указывает на общую сумму затрат на поставку продукции или оказание услуг и напрямую влияет на общую стоимость контракта.
4. Сумма в рублях объекта закупки, отражающая сумму затрат, связанных с объектом закупки. Большая сумма указывает на значимость и важность поставляемого продукта или услуги и может существенно влиять на общую стоимость контракта.

Все эти факторы напрямую или косвенно влияют на стоимостные параметры госконтракта. Их учет при прогнозировании ценности способствует оценке и планированию затрат заказчиками, а также формированию ценообразования поставщиками. Для более точного прогноза стоимости госконтракта необходимо учитывать дополнительные факторы, такие как инфляция, валютные колебания, стоимость сырья и т.д.

Программный комплекс оценки рисков исполнения государственных контрактов

Программный комплекс оценки рисков исполнения государственных контрактов представляет собой web-приложение, которое можно развернуть как на локальных серверах потенциальных заказчиков, так и в глобальной сети.

Он состоит из 3 основных модулей:

1. Оценка потенциального заказчика. В данном модуле достаточно ввести ИНН потенциального партнера и получить о нем необходимую информацию. Заказчик вводит ИНН потенциального закупщика и там появляется информация о нем, включая все налоговые, финансовые и юридические составляющие, в любом удобном виде (таблицы, графики, диаграммы и т.д.).
2. Оценка вероятности исполнения контракта.

Заполнив предлагаемые формы, мы можем показать вероятность исполнения контракта.

Отметим, что в данном модуле можно вывести наиболее важные для прогноза признаки, в любой форме представления, которые пользователь может самостоятельно анализировать.

3. Оценка прогноза стоимости и сроков выполнения контракта

Заполнив предлагаемые формы, пользователь может получить информацию о стоимости контракта и сроках его выполнения. Кроме этого, модуль может предоставить графики, диаграммы или статистические показатели, которые помогут понять точность и достоверность прогноза.

Заключение

В результате исследования получены следующие результаты:

Собран уникальный набор данных, состоящий из более 83 тысяч данных по более чем 190 признаков с систем: Реестра государственных закупок Единой информационной системы (ЕИС) (<https://zakupki.gov.ru/>); Реестра недобросовестных поставщиков (РНП) ЕИС (<https://zakupki.gov.ru/epz/dishonestsupplier/search/results.html>); Информационной системы СПАРК (<https://spark-interfax.ru/>).

Также разработаны автоматизированные системы сбора и обновления данных, которые можно развернуть на серверах потенциальных заказчиков.

На основе анализа интерпретируемых методов машинного обучения для решения задач прогноза исполнения государственных контрактов, методов анализа неструктурированной информации большого объема, определения критериев для построения моделей машинного обучения были построены интерпретируемые модели машинного обучения для решения задач:

- оценка риска выбора недобросовестного подрядчика;
- оценка риска невыполнения контракта в срок;
- оценка вероятных сроков и стоимости реализации контракта.

Модели были протестированы и метрики обучения показали высокие оценки.

Каждый из алгоритмов разработан с учетом следующих критериев: надежность, понятность и интерпретируемость, универсальность, масштабируемость и гибкость.

Интеллектуальная система, направленная на оценку риска выбора недобросовестного подрядчика, риска невыполнения контракта в срок и оценку вероятных сроков и стоимости реализации контракта, учитывает важность признаков при оценке и принятии решений. Важность признаков указывает на то, как сильно определенный признак или фактор влияет на риск или оцениваемую характеристику. Учет важности признаков позволяет системе сосредоточиться на наиболее информативных факторах и принимать во внимание их влияние на риски и оцениваемые характеристики. Это может помочь снизить шум и улучшить качество оценок и прогнозов, что в свою очередь улучшит принятие решений в контексте выбора подрядчиков, выполнения контрактов и оценки сроков и стоимости их реализации.

В результате исследования был разработан программный комплекс для интеллектуального прогнозирования выполнения государственных контрактов. Этот комплекс предоставляет возможность проведения более точного анализа рисков с

использованием методов анализа неструктурированной информации, моделей машинного обучения и интерпретируемых методов. Это позволяет повысить эффективность контроля выполнения государственных контрактов и снизить вероятность возникновения коррупции и нарушений.

Исследование доказывает значимость и применимость методов и моделей машинного обучения в области государственных контрактов. Программный комплекс для интеллектуального прогнозирования исполнения государственных контрактов предоставляет новые возможности для улучшения процессов контроля и принятия решений в сфере государственного закупочного дела.

Библиография

1. Геллер А. Я. Анализ причин расторжения контрактов в системе государственных и муниципальных закупок Российской Федерации // Вестник университета. 2022. №. 5. С. 5-12.
2. Золотухина М. М., Половникова Н.А. Риски при выборе поставщиков и заключении контрактов // Экономика и бизнес: теория и практика. 2023. №. 8 (102). С. 86-90.
3. Гендлина Ю. Б. и др. Риски строительных контрактов, заключенных с муниципальным заказчиком // Управление и экономика народного хозяйства России. 2022. С. 113-120.
4. Алейникова М. Ю., Голованов Д. А. Модели совершенствования системы внутреннего контроля осуществления государственных закупок в Российской Федерации // Управленческий учет. 2022. № 7. С. 12-19. doi:10.25806/uu7202212-19.
5. Елисеев Д. А., Романов Д. А. Машинное обучение: прогнозирование рисков госзакупок // Открытые системы. СУБД. 2018. № 2. С. 42-44.
6. Немцева Ю. В., Миронец О. Б. Управление рисками компаний отрасли информационных технологий на рынке b2g // Управленческий учет. 2022. № 12. С. 100-109.
7. Лавлинский С. М., Панин А. А., Плясунов А. В. Модель формирования экономической политики с учётом трансакционных издержек и страхования рисков разрыва контрактов // Дискретный анализ и исследование операций. 2022. № 3 (29). С. 45-63. doi:10.33048/daio.2022.29.738.
8. Скобелев В.В. Проверка конкурсантов на этапе предварительного отбора является неотъемлемым условием успешного выполнения государственных контрактов // Национальная безопасность и стратегическое планирование. 2016. № 4 (16). С. 117-119.
9. Kultin, N.B., Kultin, D.N., Bauer, R.V. (2020) Application of machine learning technology to analyze the probability of winning a tender for a project. Proceedings of the Institute for System Programming of the RAS, 32(2), 29-36.
10. Nepomnyashchy, O. M., Barzylovich, D. V., Medvedchuk, O. V. (2018) Risk-management as an instrument of the public regulation of risk assessment in construction and operation of buildings. Публічне урядування, 5(15), 188-202. doi: 10.32689/2617-2224-2018-15-5-188-202.
11. Бекетнова Ю. М. Анализ государственных закупок методами машинного обучения в целях противодействия отмыванию доходов // Экономика и предпринимательство. 2022. № 5(142). С. 144-148. doi:10.34925/EIP.2022.142.5.027.
12. Манцава А. З. Классификация завершенных государственных закупок с использованием алгоритма машинного обучения // Академическая публицистика.

2022. № 5-2. С. 72-78.

13. Устинова В.В. Минимизация рисков при заключении государственных контрактов в сфере закупок для нужд органов внутренних дел // Проблемы в российском законодательстве. 2016. № 4. С. 119-123.
14. Старосельцева М.М., Устинова В.В. Гражданско-правовые риски при заключении государственных контрактов в сфере закупок для нужд органов внутренних дел // Вестник Тверского государственного университета. Серия: Право. 2016. № 4. С. 25-36.
15. Коровина А.Н. Система мер по парированию рисков контрактов на поставку продукции по государственному оборонному заказу // Наука и военная безопасность. 2016. № 2 (5). С. 87-96.
16. Черняев Е. В., Хайтбаев В. А. Вероятностные методы оценки рисков в системе программно-целевого планирования государственного оборонного заказа с применением иерархических моделей // Прикладные экономические исследования. 2022. №. 3. С. 24-29.
17. Зембатов М.Р. Анализ факторов риска поставщика при заключении государственных контрактов // Научно-исследовательский финансовый институт. Финансовый журнал. 2017. № 6 (40). С. 75-97.
18. Иванов Н.А., Валлпетерс М.Л., Киреев И.А. "Большие данные" и машинное обучение при управлении рисками невыполнения обязательств по контрактам в строительной отрасли // Промышленное и гражданское строительство. 2019. № 5. С. 81-87.
19. Травкина Л. С., Лисин П. В., Межевикина Е. А. Экспертная оценка рисков поставки продукции, не соответствующей требованиям контракта // Контроль качества продукции. 2021. № 2. С. 24-28.
20. Асадуллаев Р. Г., Ломакин В. В., Путинцева Н. П. Разработка средств оценки проектных рисков при создании информационных систем для сферы государственных услуг // Научно-технический вестник Поволжья. 2017. № 5. С. 120-122. doi: 10.24153/2079-5920-2017-7-5-120-122.
21. Никитин П.В., Долгов В.И., Горохова Р.И., Коровин Д.И., Бахтина Е.Ю. Методика оценки важности признаков при анализе выполнения государственных контрактов // Национальная безопасность. 2023. № 4. С. 1-18. doi: 10.7256/2454-0668.2023.4.44013

Результаты процедуры рецензирования статьи

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов издательства можно ознакомиться [здесь](#).

Предмет исследования. В рецензируемой статье «Методология оценки рисков выполнения государственных контрактов с использованием средств машинного обучения» предметом исследования является методология (система приемов и способов) оценки рисков в сфере закупок и заключения государственных контрактов в различных отраслях национальной экономики, в том числе, их интеллектуальное прогнозирование посредством современных информационно-коммуникационных технологий (программного обеспечения и искусственного интеллекта).

Методология исследования. Методологический аппарат составили следующие диалектические приемы и способы научного познания: анализ, абстрагирование, индукция, дедукция, гипотеза, аналогия, синтез, типология, классификация,

систематизация и обобщение. В работе использовалось сочетание эмпирической и теоретической информации. Применение современных методов научного познания позволило изучить сложившиеся подходы, взгляды на предмет исследования, выработать авторскую позицию и аргументировать ее.

Актуальность исследования. Актуальность темы исследования, заявленной авторами, не вызывает сомнения. Цифровизация предполагает внедрение новых технологий во все отрасли экономики, что отвечает общемировым тенденциям. Авторы отмечают, что «В данной работе было проведено исследование, нацеленное на комплексный анализ и управление рисками при выполнении государственных контрактов. В рамках исследования предложены методы и подходы, основанные на применении машинного обучения и интеллектуального анализа данных, которые позволяют рассмотреть риски системно и эффективно. Данное исследование заполняет пробел в существующей литературе и предлагает новый подход к управлению рисками в данной области». Эти обстоятельства указывают на актуальность доктринальных разработок по заявленной тематике и их практическую значимость.

Научная новизна. Не подвергая сомнению важность проведенных ранее научных исследований, послуживших теоретической базой для данной работы, тем не менее, можно отметить, что в этой статье впервые сформулированы заслуживающие внимания положения, например: «...В результате исследования был разработан программный комплекс для интеллектуального прогнозирования выполнения государственных контрактов. Этот комплекс предоставляет возможность проведения более точного анализа рисков с использованием методов анализа неструктурированной информации, моделей машинного обучения и интерпретируемых методов. Это позволяет повысить эффективность контроля выполнения государственных контрактов и снизить вероятность возникновения коррупции и нарушений...». Авторами по результатам написания статьи сделан ряд теоретических выводов и предложений, что указывает не только на важность этого исследования для науки, но и определяет его практическую значимость.

Стиль, структура, содержание. Статья написана научным стилем, использована специальная терминология. Однако встречаются в тексте повторы слов в отдельных предложениях, например: «данное... в данной...». Такие предложения следует перефразировать. В целом же материал изложен последовательно и ясно. Для лучшего восприятия и наглядности информации авторами предложены рисунки и таблицы. Статья структурирована. Введение отвечает требованиям, предъявляемым к этой части научной статьи. В заключении сформулированы основные результаты, которые достигли авторы в ходе исследования. Тема раскрыта, содержание статьи соответствует ее названию.

Библиография. Авторами использовано достаточное количество доктринальных источников, представлены ссылки на публикации последних лет. Ссылки на источники оформлены в соответствии с требованиям библиографического ГОСТа.

Апелляция к оппонентам. По отдельным вопросам заявленной тематики представлена научная дискуссия, обращения к оппонентам корректные. Все заимствования оформлены ссылками на автора и источник опубликования.

Выводы, интерес читательской аудитории. Статья «Методология оценки рисков выполнения государственных контрактов с использованием средств машинного обучения» рекомендуется к опубликованию. Статья соответствует тематике журнала «Программные системы и вычислительные методы». Статья написана на актуальную тему, имеет практическую значимость и отличается научной новизной. Данная статья может представлять интерес для широкой читательской аудитории, прежде всего, специалистов в области экономики, информационных технологий, фундаментальных наук и практики программирования, а также, будет полезна для преподавателей и обучающихся экономических и технических вузов и факультетов.

