

Прогнозирование рисков неисполнения мероприятий национальных проектов

А.А. Червяков

Аннотация—Рассматривается задача мониторинга и прогнозирования рисков неисполнения мероприятий национальных и федеральных проектов в рамках Государственной автоматизированной информационной системы «Управление» (ГАСУ). Цель исследования состоит в создании технологического решения по формированию инструментов моделирования и прогнозирования неисполнения национальных и федеральных проектов на основе методов машинного обучения, обеспечивающих качественных прогноз для систем принятия решений федерального уровня. Машинное обучение использует исторические данные, хранящиеся в ГАСУ. Разработана структурно-устойчивая модель прогнозирования рисков просрочки контрольных точек на основе ансамбля методов машинного обучения. Модель использует ключевые индикаторы и сведения о контрольных точках, включая финансовые и натуральные показатели. Выполнена интеграция результатов классификации ансамблем методов машинного обучения с данными о финансовых рисках мероприятий и допущенных просрочек ранее в рассматриваемом году. Модель апробирована на данных за 2023–2024 гг. и внедрена в эксплуатацию в ГАСУ в 2025 г. Результаты моделирования показывают более высокие оценки по метрикам качества при прогнозировании на уровне мероприятий по сравнению с уровнем отдельных контрольных точек. Перспективное направление развития является расширение признакового пространства за счёт дополнительных данных (бюджетные, налоговые, методологические и иные риски) и прогнозирование уровня достижения мероприятий. Сформированное технологическое решение позволит повысить качество принятия решений, имеющих существенное значение в сфере управления национальными и федеральными проектами.

Ключевые слова— национальные проекты, федеральные проекты, контрольные точки, машинное обучение, прогнозирование рисков, ГАСУ, структурная устойчивость модели.

1. ВВЕДЕНИЕ

Статья посвящена решению задачи мониторинга хода выполнения национальных проектов (НП) и федеральных проектов (ФП) в рамках Государственной автоматизированной информационной системы «Управление» (ГАС «Управление», далее — ГАСУ)[1, 2]. ГАСУ представляет собой единую государственную информационную систему, обеспечивающую сбор, учёт, обработку и анализ данных из государственных и муниципальных информационных ресурсов, официальной статистики и иных сведений для поддержки принятия управленческих решений в сфере государственного управления.

Статья получена 10 марта 2026.

Червяков А.А. Федеральное Казначейство Российской Федерации, Москва, Россия, (e-mail: achervyakov@roskazna.ru).

Закрывающая часть портала ГАСУ предназначена для сотрудников органов государственной власти и местного самоуправления. Оценка хода и результатов выполнения НП и ФП осуществляется по контрольным точкам (КТ).

Цель исследования — создание технологического решения по формированию инструментов моделирования и прогнозирования неисполнения национальных и федеральных проектов на основе методов машинного обучения, обеспечивающих качественных прогноз для систем принятия решений федерального уровня.

Важность решение задач определится ее значением для развития цифрового пространства Российской Федерации и регулируется следующими нормативными документами:

Указ Президента РФ от 07.05.2024 № 309 «О национальных целях развития Российской Федерации на период до 2030 года и на перспективу до 2036 года»

и Постановление Правительства РФ от 25.12.2009 № 1088 «О государственной автоматизированной информационной системе „Управление“»,

Постановление Правительства РФ от 27.11.2015 № 1278 «О федеральной информационной системе стратегического планирования и внесении изменений в Положение о государственной автоматизированной информационной системе „Управление“».

ПП № 1278 от 27.11.2015 расширяет функционал системы «Управление», включая обеспечение информационного обеспечения стратегического планирования. Система взаимодействует с другими информационными системами (например, «Электронный бюджет»), что позволяет интегрировать данные для анализа и прогнозирования. В изменениях к Положению о системе «Управление» закреплено обеспечение мониторинга и контроля реализации документов стратегического планирования, что включает оценку рисков.

Дополнительные правовые основания регламентирующие работу в части модели прогнозирования рисков - временный порядок управления и мониторинга Единого плана по достижению национальных целей развития РФ (утверждён резолюцией Д. Ю. Григоренко от 25.06.2025). Этот документ определяет методологию формирования и мониторинга Единого плана, который включает национальные проекты. В нём закрепляются требования к системам сбора и обработки данных, а также необходимость использования цифровых инструментов для прогнозирования и оценки прогресса. Прогнозирование рисков неисполнения мероприятий является частью мониторинга реализации Единого плана.

Таким образом, актуальность и важность построения модели прогнозирования рисков неисполнения меро-

приятый национальных проектов соответствует указанным нормативным актам, так как: направлена на достижение национальных целей, установленных Указом Президента № 309; реализуется в рамках государственной автоматизированной информационной системы «Управление», что обеспечивает интеграцию данных и соблюдение требований к их обработке; способствует повышению эффективности государственного управления через анализ и прогнозирование, что прямо предусмотрено Положением о системе «Управление»; соответствует задачам мониторинга и контроля реализации документов стратегического планирования, закреплённым в Постановлении № 1278.

II. ИСХОДНЫЕ ДАННЫЕ И МЕТОДЫ

Исходные данные для исследования представлены обезличенными и нормированными сведениями из ГАСУ. В информационную подсистему «Федерация» ГАСУ входят:

- модуль сбора и обработки данных;
- единое хранилище данных;
- система ведомственных проектных офисов (СВПО).

В системе аккумулируются: финансовые и натуральные показатели НП и ФП; данные паспортов проектов, включая ключевые показатели и методы расчёта уровня их достижения; сведения о мероприятиях и соответствующих контрольных точках [3].

На рис. 1. показано графовое представление контрольных точек одного из НП. Каждая фиолетовая точка – это контрольная точка текущего года, жёлтые точки, агрегирующие фиолетовые, соответственно, отдельные годы реализации НП.

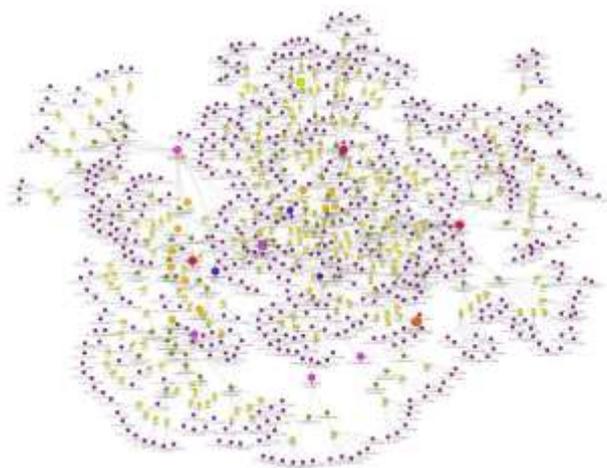


Рис. 1. Графовое представление проектов в ГАСУ

Ключевым компонентом организационно-технического управления выступает система ведомственных проектных офисов (СВПО) — подсистема инцидент-менеджмента в рамках ГАСУ[4].

В контексте рассматриваемой задачи под инцидентом понимается нарушение сроков выполнения контрольных точек (КТ) проектов. Причины возникновения инцидентов могут носить как объективный, так и субъективный характер (в т.ч. связанные с деятельностью исполнителей, ответственных за внесение данных в ин-

формационные системы).

Механизм реагирования СВПО, ответственного за реализацию проекта, предусматривает последовательное выполнение следующих действий:

1. Составление отчёта о причинах опоздания и планируемых мерах по закрытию КТ.
2. Реализация работ, необходимых для закрытия КТ.
3. При необходимости — формирование запроса на корректировку параметров НП и ФП (в части сроков и/или параметров соответствующей КТ и/или связанных проектов).

Внедрение СВПО в апреле 2022 года способствовало снижению количества просроченных КТ. До запуска организационно-технического сопровождения СВПО свыше 20% КТ выполнялись с нарушением сроков.

На рис. 2 также приведён алгоритм рассылки индикаторов в проектные офисы НП и ФП. В частности, жёлтый индикатор сигнализирует о возникновении риска просрочки КТ.

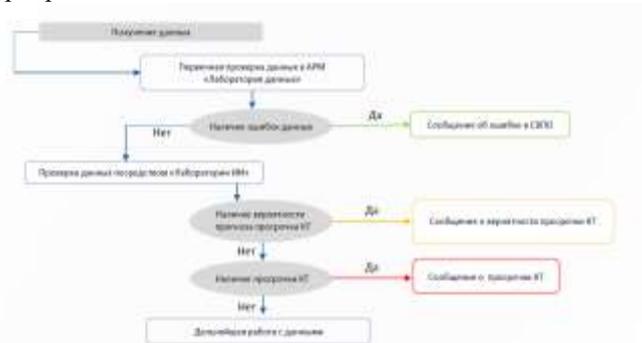


Рис. 2. Процедура опережающего оповещения о возможных задержках в сроках выполнения КТ

С учётом изложенного актуальной задачей представляется разработка интеллектуальных инструментов на базе методов машинного обучения. Такие инструменты должны анализировать данные, аккумулируемые в ГАСУ и выявлять КТ проектов с высоким риском нарушения сроков выполнения, а также своевременно предоставлять соответствующую информацию ответственным проектным офисам для превентивного принятия управленческих решений.

Для этого ведётся разработка модели, обладающей структурной устойчивостью. Определим структурную устойчивость математической модели следующим образом [4].

Пусть имеется обучающая выборка, состоящая из набора входных данных D и набора результатов R , пусть тогда на основе машинного обучения может быть построена модель M :

$$D \xrightarrow{M} R, \quad (1)$$

тогда при использовании модели на реальных (или преобразованных) данных \tilde{D}

$$\tilde{D} \xrightarrow{M} R', \quad (2)$$

выполнено –

$$\|R' - \tilde{R}\| \rightarrow \min, \quad (3)$$

где \tilde{R} – реальный (преобразованный) набор результатов, то такая модель M является структурно-устойчивой.

Структурно-устойчивая модель на тестовых данных

должна давать достоверный результат, сопоставимый с результатами на обучающей выборке. Критерием адекватности моделей могут быть использованы стандартные метрики оценки, например, точность, полнота, F1-мера, но примененные к преобразованным данным [5].

Предлагается считать нечувствительность моделей к допустимым преобразованиям данных как критерий структурной устойчивости моделей условием их применения в практике.

III. ХАРАКТЕРИСТИКА ДАННЫХ И ИСХОДНОЕ ПРИЗНАКОВОЕ ПРОСТРАНСТВО

В качестве исходных данных использован датасет с информацией о КТ за период с 2022 по 2024 год (табл. 1). После очистки данных от записей по условию, что либо отсутствует отметка о фактической дате завершения (fact_date), либо присутствует отметка о неактуальности КТ (dm1_type_result), объем датасета составил 13965 записей. По каждому году датасет содержит данные за полный календарный год.

Таблица 1. Количественные характеристики исходных данных

Год	КТ, всего	КТ, с реализованным риском просрочки	КТ, своевременно выполненные	Мероприятия
2022	4237	749	3488	617
2023	4366	466	3900	633
2024	5362	463	4899	784

Каждая КТ, помимо отметки о наличии/отсутствии опоздания, после процедуры генерации признаков характеризуется 14 категориальными, 3 из которых – бинарными, и 4 целочисленными признаками [6]. В признаковое пространство не включались дублирующие поля и поля, значения которых нельзя применить для задачи машинного обучения. В табл. 2 приведено описание признаков.

Таблица 2. Признаковое пространство

Описание признака	Тип данных признака	Количество категорий в признаке
ID НП	Номинальный	16
ID ФП	Номинальный	85
ФОИВ, ответственный за НП	Номинальный	13
ФОИВ, ответственный за ФП	Номинальный	21
Наименование типа результата	Номинальный	22
Тип КТ	Номинальный	5
Наименование типа КТ	Номинальный	99
Дата завершения работ по КТ (месяц)	Номинальный	12
Дата завершения работ по КТ (день в месяце)	Номинальный	31
Дата завершения работ по КТ (день недели)	Номинальный	7
ОГРН ФОИВ	Номинальный	37
Денежный результат	Бинарный	2
Реализуется регионами	Бинарный	2
Финансовое обеспечение ФП	Бинарный	2

На рис. 3–5 показано распределение КТ по НП и количество опозданий в соответствующих категориях. Видно, что существуют НП, по которым после 2022 года практически не наблюдается просрочек, в то время как по другим доля просрочки присутствует, хотя и со-

кращается.

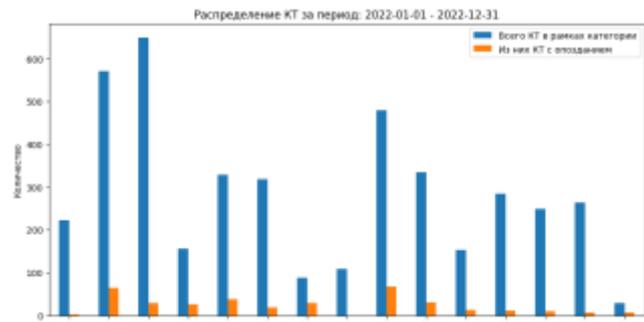


Рис. 3. Распределение КТ в рамках признака «pr_code» по данным за 2022 год

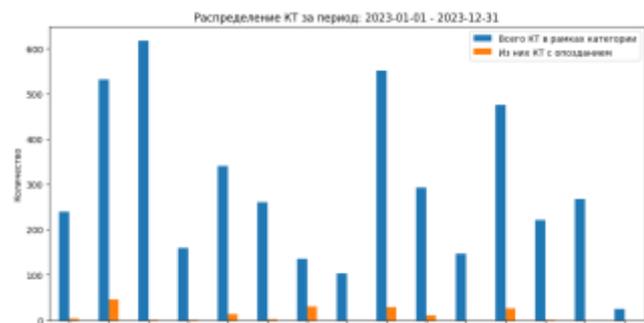


Рис. 4. Распределение КТ в рамках признака «pr_code» по данным за 2023 год

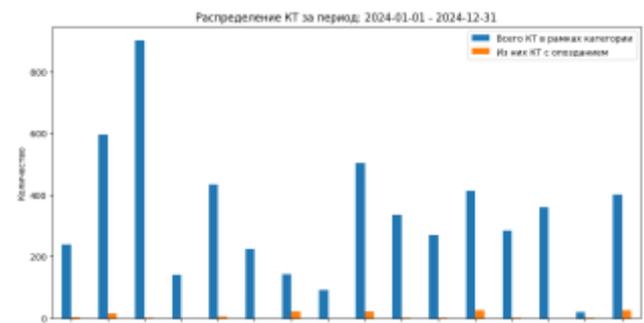


Рис. 5. Распределение КТ в рамках признака «pr_code» по данным за 2024 год

Проведённый анализ для прочих признаков подтверждает наличие устойчивых закономерностей в рассматриваемом признаковом пространстве, пригодных для построения прогностических моделей. Кроме того, оценка динамики этих закономерностей позволяет определить период их дрейфа — временной интервал, в течение которого статистические характеристики и взаимосвязи между признаками сохраняют достаточную стабильность для корректного прогнозирования.

IV. РЕЗУЛЬТАТЫ ПРОГНОЗИРОВАНИЯ РИСКОВ ПРОСРОЧКИ КТ С ПОМОЩЬЮ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Перед использованием для машинного обучения данные проходят предобработку [7]. Записи о КТ из большего класса (КТ без просрочки), идентичные по 11 и более категориальным признакам записям из меньшего класса (КТ с просрочкой), удаляются. Затем производится оверсэмплинг методом синтезированных меньшинств для номинальных признаков (SMOTEN).

С помощью подготовленных данных производилось

обучение ансамбля моделей с мягким голосованием, выполняющего классификацию КТ по двум классам: с риском и без риска опоздания. В него входят классификаторы на основе дерева принятия решений, метода опорных векторов, случайного леса, К ближайших соседей, логистической регрессии, адаптивного бустинга и линейный классификатор с обучением на основе стохастического градиентного спуска.

В результате применения ансамбля получены прогнозы рисков просрочки по КТ. Для прогноза на 2023 год в качестве обучающей выборки использованы данные 2022 года, для 2024 года – данные за 2022 и 2023 годы. Результат по 2023 году (табл. 3 и рис. 6) показал точность равную 0.17, полноту 0.78 и F1-меру равную 0.28. Результат по 2024 году (табл. 3 и рис. 7) демонстрирует точность равную 0.15, полноту 0.71 и F1-меру равную 0.25. Это говорит о том, что большая часть КТ с опозданием успешно выявляется. При этом 33.76% КТ за 2023 год и 33.7% КТ за 2024 год были отмечены как имеющие риск опоздания, хотя были завершены вовремя.

Таблица 3. Оценки эффективности модели при выявлении риска просрочки КТ по 14 признакам

Метрика	Значение	
	По данным за 2023 год	По данным за 2024 год
Точность (precision)	0.17	0.15
Полнота (recall)	0.73	0.71
F1-мера	0.28	0.25
Доля выявленной просрочки в днях	0.69	0.64

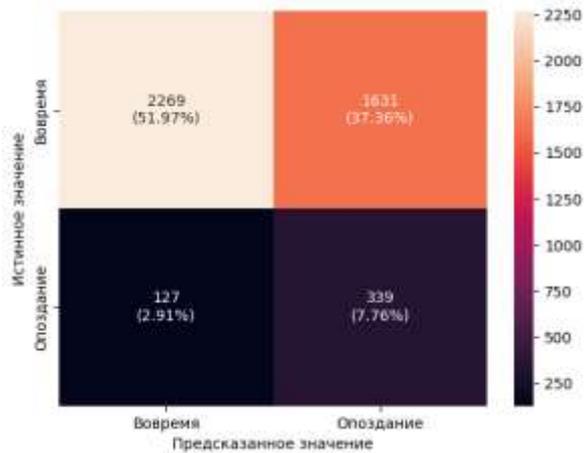


Рис. 6. Матрица ошибок при прогнозировании риска просрочки выполнения КТ НП за 2023 год по 14 признакам

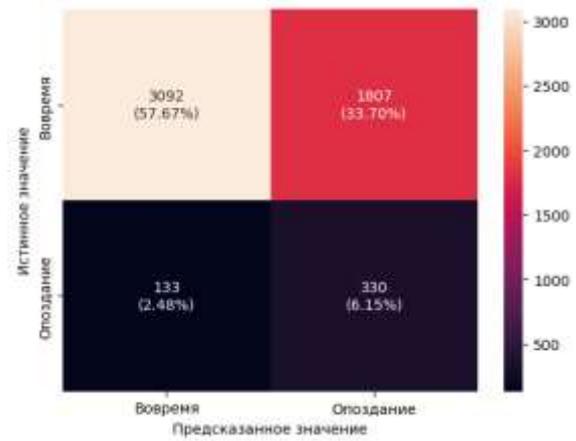


Рис. 7. Матрица ошибок при прогнозировании риска просрочки выполнения КТ НП за 2024 год по 14 признакам

Далее признаковое пространство было дополнено путем генерации 4 новых признаков. В него вошли целочисленные признаки, отражающие количество КТ в рамках мероприятия, количество предшествующих и последующих КТ в рамках мероприятия, а также КТ в рамках мероприятия, имеющие установленную дату завершения работ в том же месяце, что и рассматриваемая КТ. После дополнения признакового пространства были проведены оценки качества классификации по данным за 2023 и 2024 год аналогично экспериментам, описанным выше.

Дополнение признакового пространства позволило повысить качество классификации, при этом не привело к значимому увеличению числа ложных срабатываний.

V. МОДЕЛЬ ПРОГНОЗИРОВАНИЯ РИСКОВ НЕИСПОЛНЕНИЯ МЕРОПРИЯТИЙ ПО ДАННЫМ КТ

Пусть в году дано множество мероприятий НП

$R = \{X_i, i \in \overline{1, n}\}$, где n – количество мероприятий НП.

Каждое мероприятие – это множество контрольных точек

$X_i = \{\mathbf{x}_i^j, j \in \overline{1, m_i}\}$, где m_i – количество КТ в i -м мероприятии НП, а \mathbf{x}_i^j – вектор признаков КТ. Наличие

просрочки для КТ определяется как:

$$f(\mathbf{x}_i^j) = \begin{cases} 1 & \text{если просрочка допущена,} \\ 0 & \text{в иных случаях.} \end{cases} \quad (4)$$

Каждая КТ характеризуется вектором признаков, таких как НП, ФП, ответственный ФОИВ, реализация регионами, назначенная дата завершения работ и другими.

Реализована модель на основе ансамбля методов машинного обучения, выполняющая классификацию КТ по наличию риска просрочки, с использованием вектора признаков КТ \mathbf{x}_i^j , такая, что:

$$f'(\mathbf{x}_i^j) = \begin{cases} 1 & \text{если спрогнозирован риск просрочки,} \\ 0 & \text{в иных случаях.} \end{cases} \quad (5)$$

В ансамбль входят классификаторы на основе дерева решений, опорных векторов, случайного леса, k-ближайших соседей, логистической регрессии, адаптивного бустинга и линейный классификатор с обучением методом стохастического градиентного спуска. Итого-

вый ответ формируется с использованием мягкого голо-
сования.

Для того, чтобы перейти от уровня отдельных КТ к уровню мероприятий НП необходимо дать определение мероприятию с реализованным риском просрочки. Мероприятием НП с реализованным риском просрочки назовем такое, что:

$$g(X_i) = f(x_i^1) \vee f(x_i^2) \vee \dots \vee f(x_i^{m_i-1}) \vee f(x_i^{m_i}). \quad (6)$$

Иными словами, это такое мероприятие, для которого одна или более КТ были выполнены с просрочкой.

Тогда мероприятие, для которого прогнозируется риск просрочки, будет:

$$g'(X_i) = f'(x_i^1) \vee f'(x_i^2) \vee \dots \vee f'(x_i^{m_i-1}) \vee f'(x_i^{m_i}). \quad (7)$$

Т. е. это мероприятие НП, в рамках которого одна или более КТ были отмечены ансамблем как имеющие риск просрочки.

Для оценки эффективности модели проведено 2 эксперимента. В первом в качестве обучающей выборки используются данные 2022 года, а оценка выполняется по данным 2023 года (рис. 8). Во втором для обучения взяты данные за 2022 и 2023 годы, а оценка выполняется по данным 2024 года (рис. 9). Полученные данные о просрочке КТ (4) и результатах классификации (5) были объединены на уровне результатов НП с использованием (6) и (7).

Оценки эффективности при выявлении риска просрочки на уровне мероприятий НП приведены в табл. 4. Большая часть мероприятий НП с просрочкой была выявлена, а полученные оценки на данных за 2023 и 2024 годы находятся на близком уровне. Модель, обученная на данных 2022 и 2023 годов, показывает точность равную 0.39 и полноту 0.81 при проверке на данных 2024 года.

Таким образом, данные КТ и результатов их классификации были объединены по мероприятиям, в результате чего отмечаются более высокие оценки по метрикам качества классификации, чем при прогнозировании на уровне отдельных КТ.

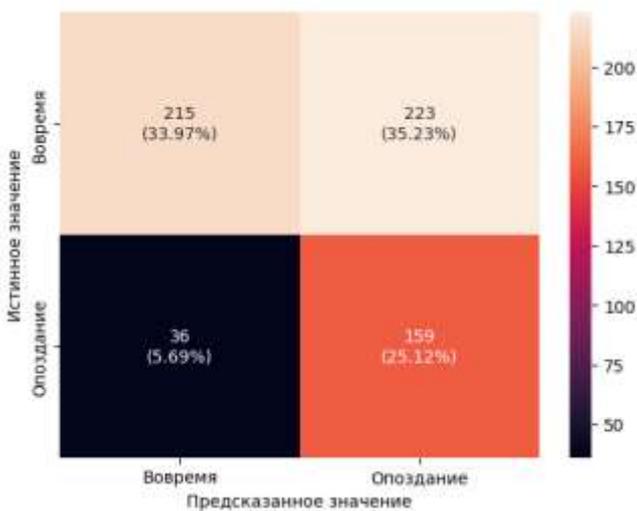


Рис. 8. Матрица ошибок при прогнозировании риска просрочки по мероприятиям НП за 2023 год по 18 признакам

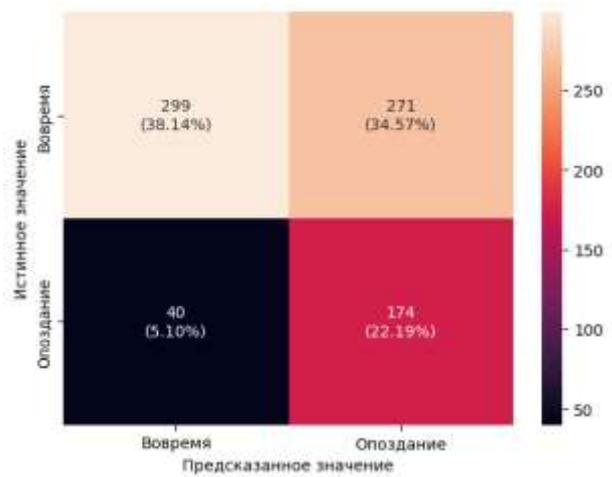


Рис. 9. Матрица ошибок при прогнозировании риска просрочки по мероприятиям НП за 2024 год по 18 признакам

Таблица 4. Оценки эффективности модели при выявлении риска просрочки на уровне мероприятий НП

Метрика	Значение	
	По данным за 2023 год	По данным за 2024 год
Точность (precision)	0.42	0.39
Полнота (recall)	0.82	0.81
F1-мера	0.55	0.53
Доля выявленной просрочки в днях	0.91	0.94

VI. ИНТЕГРАЦИЯ ДАННЫХ РИСКОВ МЕРОПРИЯТИЙ НП, ФИНАНСОВЫХ РИСКОВ И ДАННЫХ О РЕАЛИЗОВАННЫХ ПРОСРОЧКАХ КТ

Была рассмотрена возможность улучшения полученных результатов с помощью интеграции данных [8] из нескольких источников. В качестве дополнительного источника для выявления рисков просрочки мероприятий НП использованы данные о финансовых рисках, поступающие в течение года. Они включают сведения об инцидентах в части бюджетных обязательств (БО), кассового исполнения и лимитов бюджетных обязательств (ЛБО). Помимо этого, они включают дату формирования инцидента и идентификатор мероприятия, что позволяет объединить данные. Наличие финансового риска рассматривается как предиктор просрочки по мероприятию.

Так как данные о финансовых рисках меняются ежемесячно, прогноз с их использованием будет повторяться на протяжении текущего года для уточнения рисков на оставшийся период в году. С учетом этого в качестве предиктора просрочки по мероприятию также могут быть включены данные о ранее просроченных КТ в текущем году.

Производимая интеграция данных иллюстрируется диаграммой Венна (рис. 10). Так мероприятие, имеющее риск просрочки, это такое мероприятие, для которого присутствует либо положительное предсказание ансамбля моделей классификации, либо финансовые риски, либо наличие просрочки по КТ ранее в текущем году.



Рис. 10. Иллюстрация подхода к интеграции данных для выявления рисков просрочки

В результате выполнения прогноза на начало каждого месяца в 2024 году были получены оценки качества классификации (табл. 5). Так как данные о финансовых

Таблица 5. Оценки качества классификации при выявлении рисков просрочки в 2024 году

Мес.	Точность		Полнота		F1-мера		Доля выявленной просрочки	
	Данные анс. моделей	Интегр. данные	Данные анс. моделей	Интегр. данные	Данные анс. моделей	Интегр. данные	Данные анс. моделей	Интегр. данные
1	0.39	0.39	0.81	0.81	0.53	0.53	0.94	0.94
2	0.37	0.37	0.82	0.82	0.51	0.51	0.96	0.96
3	0.36	0.36	0.84	0.86	0.51	0.50	0.96	0.97
4	0.35	0.34	0.80	0.83	0.49	0.48	0.94	0.95
5	0.31	0.30	0.78	0.80	0.45	0.44	0.93	0.94
6	0.30	0.29	0.78	0.81	0.43	0.43	0.95	0.96
7	0.28	0.27	0.77	0.82	0.41	0.41	0.95	0.96
8	0.24	0.22	0.74	0.77	0.36	0.34	0.95	0.95
9	0.22	0.19	0.73	0.81	0.34	0.31	0.94	0.95
10	0.20	0.20	0.76	0.82	0.32	0.32	0.93	0.93
11	0.24	0.21	0.72	0.81	0.36	0.34	0.92	0.93
12	0.18	0.16	0.68	0.80	0.29	0.27	0.89	0.91
Средн.	0.287	0.275	0.769	0.813	0.417	0.407	0.938	0.946

С каждым мероприятием ассоциируется количество дней, на которые были просрочены входящие в него КТ. В результате суммирования количества дней просрочки по всем КТ и вычисления суммы количества дней просрочки по КТ в выявленных мероприятиях с риском просрочки была определена доля выявленной просрочки. Ансамбль моделей классификации позволил выявить 93.8% от общего числа дней просрочки. В результате интеграции данных пока за тель был улучшен до 94.6%. Полученные оценки классификации, позволяют заключить, что интеграция данных повысила полноту и долю выявленной просрочки при небольшом снижении точности классификации.

VII. КОНФИГУРИРОВАНИЕ МОДЕЛИ И ПРИМЕНЕНИЕ К НОВЫМ ДАННЫМ

Построенная модель включает настройки, позволяющие управлять чувствительностью [9–11] оценки риска опозданий в сроках выполнения. Так, например, можно изменить настройки просрочки на этапах обучения и классификации, тем самым настроив модель на выявление наиболее значимых опозданий, превышающих заданный порог.

На рис. 11 приведен результат классификации КТ по данным за 2024 год, без интеграции данных с данными о

финансовых рисках.



Рис. 11. Матрица ошибок после изменения настроек модели на выявление КТ с опозданием более 30 дней

Аналогично, данные объединяются в рамках мероприятия. На рис. 12 приведен результат прогнозирования риска просрочки по мероприятиям НП за 2024 год, без интеграции данных с данными о финансовых рисках.

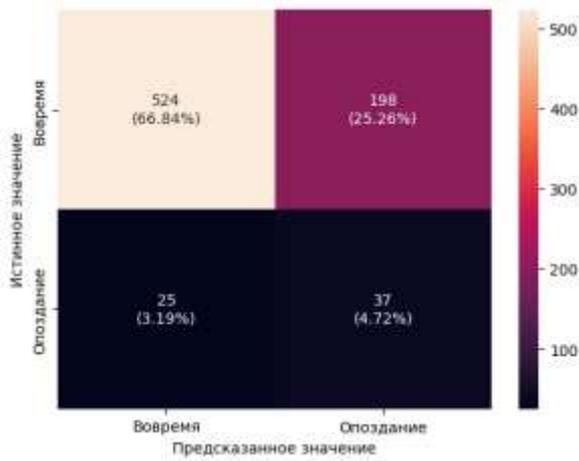


Рис. 12 Матрица ошибок после изменения настроек модели на выявление КТ с опозданием более 30 дней

С учетом указанной конфигурации была решена задача проверки функционирования модели на данных 2025 года. Он подразумевает смену перечня НП, поэтому часть НП, которая не имела предшествующих аналогов, не была включена в прогноз. Оценка и прогноз выполнялись не по всем данным, а только по данным за пройденный период 2025 года и с учетом рассматриваемого срока опоздания, превышающего 30 дней. Были взяты данные с 01.01.2025 по 25.07.2025 (рис 13).

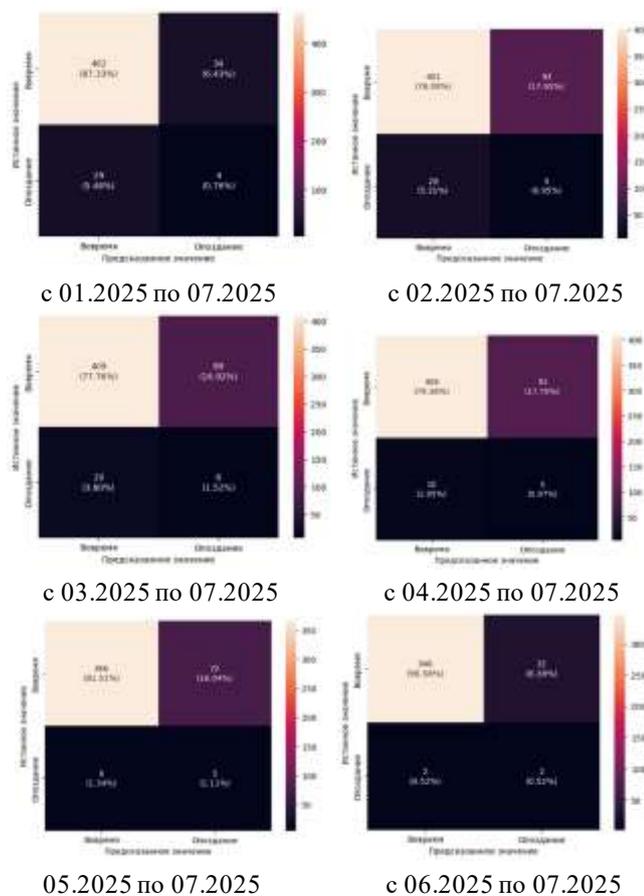


Рис. 13. Матрица ошибок для модели при прогнозировании риска просрочки по мероприятиям НП

VIII. ЗАКЛЮЧЕНИЕ

Приведено описание задачи прогноза рисков неисполнения мероприятий Национальных проектов и дано определение структурной устойчивости математической модели.

Приведена характеристика рассматриваемых данных и исходное признаковое пространство. Показаны распределения КТ по НП, ФП и другим значимым признакам. С использованием ансамбля методов машинного обучения выполнено прогнозирование рисков просрочки КТ на 2023 и 2024 год, что показало работоспособность исходной модели.

Разработана модель прогнозирования рисков неисполнения мероприятий по данным КТ. Она позволила выполнять прогноз риска просрочки на уровне мероприятий, в результате чего были получены более высокие оценки по метрикам качества классификации, чем при прогнозировании на уровне отдельных КТ.

Выполнена интеграция результатов классификации ансамблем методов машинного обучения с данными о финансовых рисках мероприятий и допущенных просрочек ранее в рассматриваемом году. Результаты интеграции показали более полное выявление просрочек по сравнению с использованием ансамбля в отдельности.

Модель внедрена в промышленную эксплуатацию в ГАСУ в 2025 году.

В дальнейших исследованиях планируется существенное расширение признакового пространства за счет дополнительной аналитики – бюджетные, налоговые, методологические, новостные и т.д. риски. А также прогноз не только факта срыва сроков мероприятия, но и уровня достижения мероприятия.

БИБЛИОГРАФИЯ

- [1] ГАСУ «Управление» [Электронный ресурс] URL: <http://gasu.gov.ru/>.
- [2] Албычев А.С. ИТ-приоритеты Казначейства России // Журнал Бюджет. 2023. № 6 (246). С. 42–47.
- [3] Албычев А.С., Червяков А.А., Никульчев Е.В., Ильин Д.Ю., Газанова Н.Ш. Разработка интеллектуальных инструментов управления организационными системами для контроля сроков в выполнении Национальных и Федеральных проектов // Теория активных систем – 55 лет (ТАС-55): труды юб. науч.-практ. конф., 18 ноября 2024. М.: ИПУ РАН, 2024. С. 11–15.
- [4] Червяков А.А., Никульчев Е.В. Разработка структурно-устойчивых моделей крупномасштабных систем управления финансами // Управление развитием крупномасштабных систем (MLSD'2025). Труды 18-й междунар. конф. М.: ИПУ РАН, 2025. С. 323–328.
- [5] Obi J.C. A comparative study of several classification metrics and their performances on data // World Journal of Advanced Engineering Technology and Sciences. 2023. V. 8. N. 1. P. 308–314.
- [6] Mukherjee M., Khushi M. SMOTE-ENC: A novel SMOTE-based method to generate synthetic data for nominal and continuous features // Applied system innovation. 2021. V. 4. N. 1. P. 18.
- [7] Никульчев Е.В., Ильин Д.Ю., Духовенский С.Е., Газанова Н.Ш., Червяков А.А. Методика оценки влияния качества данных на результативность моделей машинного обучения для определения опозданий исполнения контрольных точек проекта // International Journal of Open Information Technologies. 2025. Т. 13. № 10. С. 90–95.
- [8] Аньш Х. Применение инструментов машинного обучения и интеллектуальный анализ данных в отношении баз данных с небольшим количеством записей // Advanced Engineering Research. 2021. Т. 21. № 4. С. 346–363.
- [9] Aljohani A. Predictive analytics and machine learning for real-time supply chain risk mitigation and agility // Sustainability. 2023. V. 15. N. 20. P. 15088.

- [10] Албычев А.С., Червяков А.А., Никульчев Е.В., Ильин Д.Ю., Газанова Н.Ш. Выявление рисков просрочки достижения результатов национальных и федеральных проектов с использованием методов машинного обучения // *Фундаментальные, поисковые, прикладные исследования и инновационные проекты: сб. трудов Национальной научно-практической конференции*. – М.: АВиС ВВИА им. проф. Жуковского, 2025. С. 181–183.
- [11] Albychev, A.; Chervyakov, A.; Gazanova, N.; Ilin, D.; Nikulchev, E. Machine Learning Methods for Deadline Missing Prediction Using National Project Checkpoint Data // *Communications in Computer and Information Science*. 2026. Vol. 2604. P. 250-258. https://doi.org/10.1007/978-3-032-04761-8_19

Forecasting the risks of non-implementation of national project activities

A.A. Chervyakov

Abstract — The article addresses the problem of monitoring and forecasting risks of non-fulfilment of activities within national projects (NP) and federal projects (FP) in the framework of the State Automated Information System «Management» (GASU). The aim of the study is to create a technological solution for the formation of tools for modeling and forecasting the failure of national and federal projects based on machine learning methods, providing a high-quality forecast for federal decision-making systems. A structurally stable risk forecasting model for delays in milestone achievement (MA) has been developed. The model is based on an ensemble of machine learning methods and leverages key indicators and data on milestones, including financial metrics and physical (natural) performance indicators. The results of the classification using an ensemble of machine learning methods were integrated with data on the financial risks of events and delays admitted earlier in the year under review. The proposed model was validated using real-world data for 2023–2024 and subsequently deployed into industrial operation within the GASU in 2025. Experimental results demonstrate superior performance in terms of quality metrics when forecasting at the activity level, as compared to the level of individual milestones. A promising direction for further development involves expanding the feature space through additional analytics — incorporating budgetary, tax-related, methodological, and other types of risks — as well as enhancing forecasting capabilities to estimate the level of activity accomplishment. The developed technological solution will improve the quality of decision-making that is of significant importance in the area of managing national and federal projects.

Keywords: national projects, federal projects, milestones, machine learning, risk forecasting, GASU, structural stability of the model.

REFERENCES

- [1] GAS «Upravlenie» [Elektronnyj resurs] URL: <http://gasu.gov.ru/>.
- [2] Albychev A.S. IT-prioritety Kaznachejstva Rossii, Zhurnal Byudzhet, no. 6 (246), pp. 42–47, 2023. [in Rus]
- [3] Albychev A.S., Chervyakov A.A., Nikulchev E.V., Ilin D.YU., Gazanova N.SH. Razrabotka intellektual'nyh instrumentov upravleniya organizacionnymi sistemami dlya kontrolya srokov vypolneniya Nacional'nyh i Federal'nyh proektov, In Proc. Teoriya aktivnyh sistem – 55 let (TAS-55), 18 noyabrya 2024. M.: IPU RAN, pp. 11–15, 2024. [in Rus]
- [4] Chervyakov A.A., Nikulchev E.V. Razrabotka strukturno-ustojchivyh modelej krupnomasshtabnyh sistem upravleniya finansami, In Proc. Upravlenie razvitiem krupnomasshtabnyh sistem (MLSD'2025), 18 conf. IPU RAN, pp. 323–328, 2025. [in Rus]
- [5] Obi J.C. A comparative study of several classification metrics and their performances on data, World Journal of Advanced Engineering Technology and Sciences, vol. 8, no. 1, pp. 308–314, 2023.
- [6] Mukherjee M., Khushi M. SMOTE-ENC: A novel SMOTE-based method to generate synthetic data for nominal and continuous features, Applied system innovation, vol. 4, no. 1, p. 18, 2021.
- [7] Nikulchev E.V., Ilin D.YU., Duhovenskij S.E., Gazanova N.SH., Chervyakov A.A. Metodika ocenki vliyaniya kachestva dan'nyh na rezul'tativnost' modelej mashinnogo obucheniya dlya opredeleniya opozdaniy ispolneniya kontrol'nyh toчек proektov, International Journal of Open Information Technologies, vol. 13, no. 10, pp. 90–95, 2025. [In Rus]
- [8] Anysh H. Primenenie instrumentov mashinnogo obucheniya i intellektual'nyj analiz dannyh v otnoshenii baz dannyh s nebol'shim kolichestvom zapisej, Advanced Engineering Research, vol. 21, no. 4, pp. 346–363, 2021.
- [9] Aljohani A. Predictive analytics and machine learning for real-time supply chain risk mitigation and agility, Sustainability, vol. 15, no. 20, p. 15088, 2023.
- [10] Albychev A.S., Chervyakov A.A., Nikulchev E.V., Ilin D.YU., Gazanova N.SH. Vyyavlenie riskov prosrochki dostizheniya re-zul'tatov nacional'nyh i federal'nyh proektov s ispol'zovaniem metodov mashinnogo obucheniya. In Proc. Fundamental'nye, poiskovyje, prikladnye issledovaniya i innovacionnye proekty: sb. conf, pp. 181–183, 2025. [In Rus]
- [11] Albychev, A.; Chervyakov, A.; Gazanova, N.; Ilin, D.; Nikulchev, E. Machine Learning Methods for Deadline Missing Prediction Using National Project Checkpoint Data, Communications in Computer and Information Science, vol. 2604, pp. 250–258, 2026. https://doi.org/10.1007/978-3-032-04761-8_19

About Author

Alexander A. Chervyakov, Ph.D., Head of Department, Federal Treasury of the Russian Federation, Moscow, Russia, achervyakov@roskazna.ru, ORCID - 0000-0002-5638-8361