

## 1.5. УПРАВЛЕНИЕ РИСКАМИ С ПРИМЕНЕНИЕМ СОВРЕМЕННЫХ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И АНАЛИЗА БОЛЬШИХ ДАННЫХ

Гоглев Н.Н., Касаткина Е.В., к. ф.-м.н., Мигалин С.А., Муштак О.И. Ижевск

*Высокий уровень неопределенности современного мира в сочетании с активным развитием технологий искусственного интеллекта и анализа больших данных указывают на актуальность и своевременность применения этих технологий для повышения эффективности процессов в таком аспекте государственного и корпоративного управления, как риск-менеджмент. Внедрение подходов к управлению рисками, основанных на применении современных технологий, сегодня становится фактически одним из ключевых условий развития конкурентоспособности в различных отраслях экономики и сферах государственного регулирования. В настоящей статье рассмотрены процессы и задачи управления рисками, установленные международными и российскими стандартами риск-менеджмента. Сформулированы подходы, позволяющие, в отличие от классических технологий, более эффективно решать основные стандартные задачи риск-менеджмента за счет применения современных технологий искусственного интеллекта и анализа больших данных. Поскольку на текущий момент слабо проработаны методы идентификации новых типов рисков, авторами статьи разработан новый комплексный подход к автоматизированной идентификации новых типов рисков, основанный на применении интеллектуальных методов поиска аномалий в исторических массивах данных о деятельности организации. Рассмотрен подход к оценке вероятности наступления известных рисков и величины вероятных последствий при их наступлении, основанный на применении интеллектуальных классификационных и регрессионных моделей. Сформулированы возможные варианты практического применения интеллектуальных систем, основанных на онтологии предметной области, для повышения эффективности и качества систематизации разнородной информации, относящейся к процессам риск-менеджмента, в том числе за счет сочетания онтологических средств с моделями искусственного интеллекта.*

### Введение

В современном мире, характеризующимся высоким уровнем неопределенности, для преодоления актуальных кризисов и обеспечения устойчивого развития крайне важным является внедрение эффективных риск-ориентированных подходов к корпоративному и государственному управлению. В то же время активное развитие технологий искусственного интеллекта и анализа больших данных открывает для государства и бизнеса новые возможности оптимизации операционной и управленческой деятельности за счет цифровизации отдельных процессов и целых отраслей<sup>1</sup>.

Поэтому актуальным и своевременным является рассмотрение возможностей применения технологий искусственного интеллекта и «big data» к такой области управления, как риск-менеджмент. Данный тезис созвучен мировым тенденциям. Так, например, в актуальную редакцию документа «Концептуальные основы управления рисками организаций»<sup>2</sup> от 2017 г., являющегося одним из наиболее распространенных в профессиональном сообществе сводов лучших практик в области управления рисками, включены рекомендации по возможным аспектам применения искусственного интеллекта. В 2019 г. Федерация Европейских ассоциаций риск-менеджмента опубликовала достаточно объемную статью о возможных прикладных аспектах применения искусственного интеллекта в процессах управления рисками<sup>3</sup>. Многие государства на сегодняшний день уже сформулировали свои национальные стратегии развития искусственного интеллекта<sup>4,5,6,7</sup> и включили в них применение цифровых технологий для задач, связанных с управлением рисками в экономике и финансах, здравоохранении, государственном регулировании

<sup>1</sup> Национальная программа «Цифровая экономика Российской Федерации», утвержденная протоколом заседания президиума Совета при Президенте Российской Федерации по стратегическому развитию и национальным проектам от 4 июня 2019 г. № 7

<sup>2</sup> Committee of Sponsoring Organizations of the Treadway Commission. Enterprise Risk Management. Integrating with Strategy and Performance. URL: <https://www.coso.org/documents/2017-coso-erm-integrating-with-strategy-and-performance-executive-summary.pdf>. Дата обращения: 25.02.2022.

<sup>3</sup> Artificial Intelligence (AI) Applied to Risk Management. URL: <https://www.ferma.eu/publication/artificial-intelligence-ai-applied-to-risk-management/>. Дата обращения: 25.02.2022.

<sup>4</sup> Национальная стратегия Сингапура в области развития искусственного интеллекта. URL: <https://www.smartnation.gov.sg/files/publications/national-ai-strategy.pdf>. Дата обращения: 25.02.2022.

<sup>5</sup> Национальная стратегия Республики Корея в области развития искусственного интеллекта. URL: <https://www.msit.go.kr/bbs/view.do?sCode=eng&mId=10&mPid=9&bbsSeqNo=46&nttSeqNo=9>. Дата обращения: 25.02.2022.

<sup>6</sup> Обзор национальной стратегии Индии в области развития искусственного интеллекта. URL: <https://india-ai.gov.in/documents/pdf/NationalStrategy-for-AI-Discussion-Paper.pdf>. Дата обращения: 25.02.2022.

<sup>7</sup> Национальная стратегия Федеративной Республики Германия в области развития искусственного интеллекта. URL: [https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung\\_KI-Strategie\\_engl.pdf](https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf). Дата обращения: 25.02.2022.

и охране правопорядка, противодействии чрезвычайным ситуациям, охране окружающей среды и климата, а также в ряде других областей.

Российская национальная стратегия в этой сфере<sup>8</sup> явно закрепляет прогнозирование рисков и неблагоприятных событий в качестве одного из аспектов использования технологий искусственного интеллекта, направленных на повышение эффективности и создание принципиально новых направлений деятельности хозяйствующих субъектов в различных отраслях экономики.

Очевидно, что искусственный интеллект и технологии обработки больших данных – уже практически неотъемлемые элементы очередного витка эволюции риск-менеджмента.

В настоящей статье приводится обзор стандартных процессов и задач управления рисками, рассматриваются подходы к реализации данных задач, основанные на современных методах искусственного интеллекта и анализа больших данных, а также рассматривается ряд примеров практического применения данных подходов.

### Процессы и задачи в области управления рисками

Риск-менеджмент – достаточно зрелая дисциплина. Существуют как национальные стандарты по управлению рисками<sup>9,10,11</sup>, так и международные стандарты «ISO Guide 73» (определяет терминологический аппарат риск-менеджмента)<sup>12</sup>, «ISO 31000» (содержит принципы и рекомендации в области управления рисками)<sup>13</sup>, «ISO 31010» (описывает технологии оценки риска)<sup>14</sup>. Поскольку действующие российские стандарты<sup>15,16,17</sup> идентичны упомянутым выше международным, далее по тексту обозначенная проблематика данной статьи рассматривается на базисе именно этих стандартов.

На процессном уровне в риск-менеджменте выделяются следующие аспекты:

1) обмен информацией и консультирование – группа процессов, направленных на повышение осведомленности заинтересованных сторон и понимание ими риска для поддержки принятия решений;

2) определение области применения, среды и критериев – процессы адаптации риск-менеджмента к специфике организации;

3) оценка риска – систематические итеративные процессы идентификации, анализа и сравнительной оценки рисков;

4) обработка риска – процессы определения и выбора возможных вариантов обработки рисков, а также принятия связанных с этим управленческих решений;

5) мониторинг и пересмотр – непрерывные процессы повышения качества и эффективности риск-менеджмента;

6) документирование и отчетность – поддерживающие процессы, заключающиеся в формализации риск-менеджмента и результатов его применения.

Стандарты оценки рисков<sup>14,17</sup> содержат рекомендации по использованию технологий и методов для каждого из перечисленных аспектов риск-менеджмента. Рекомендуемые методы опираются преимущественно на экспертную оценку, а значит, требуют значительных трудозатрат на реализацию и зависят от человеческого фактора, что в свою очередь создает производные риски внесения систематических дефектов в механизмы риск-менеджмента, обусловленные ошибочными или необъективными мнениями задействованных экспертов. Стоит также отметить, что крайне важная роль в успешном применении предлагаемых классических технологий отводится процессам сбора из разнородных источников, обработки, структурирования и передачи больших объемов информации, а также систематизации знаний, возникающих по результатам риск-менеджмента.

Из приведенных выше результатов рассмотрения стандартов оценки рисков<sup>14,17</sup> вытекают некоторые идеи относительно совершенствования процессов риск-менеджмента. Во-первых, целесообразной и очевидной видится возможность применения математических моделей искусственного интеллекта в ключевых задачах оценки рисков, таких как идентификация и выявление новых типов рисков, а также оценка вероятности наступления рисков и величины вероятных последствий при их наступлении. Во-

<sup>8</sup> Национальная стратегия развития искусственного интеллекта на период до 2030 года, утвержденная Указом Президента Российской Федерации от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации»

<sup>9</sup> Federation of European Risk Management Associations. A Risk Management Standard. URL: <https://www.ferma.eu/app/uploads/2011/11/a-risk-management-standard-english-version.pdf>. Дата обращения: 25.02.2022.

<sup>10</sup> CAN/CSA ISO 31000 Risk Management – Principles and Guidelines

<sup>11</sup> JSA - JIS Q 2001. Guidelines for Development and Implementation of Risk Management System

<sup>12</sup> ISO Guide 73:2009 «Risk management - Vocabulary - Guidelines for use in standards»

<sup>13</sup> ISO 31000:2018 «Risk management – Guidelines»

<sup>14</sup> IEC 31010:2019 «Risk management - Risk assessment techniques», NEQ

<sup>15</sup> ГОСТ Р 51897–2011. «Национальный стандарт Российской Федерации. Менеджмент риска. Термины и определения»

<sup>16</sup> ГОСТ Р ИСО 31000–2019 «Национальный стандарт Российской Федерации. Менеджмент риска. Принципы и руководство»

<sup>17</sup> ГОСТ Р 58771–2019. «Национальный стандарт Российской Федерации. Менеджмент риска. Технологии оценки риска»

вторых, практически во все процессы управления рисками могут быть встроены интеллектуальные информационные системы, основанные на онтологии предметной области. Внедрение таких систем позволяет создать условия повышения эффективности и качества формализации знаний в области применения риск-менеджмента.

#### **Оценка рисков с применением моделей искусственного интеллекта**

##### ***Идентификация новых типов рисков. (знаки препинания!)***

Идентификация рисков – это итеративный процесс поиска новых типов рисков и профилирования их основных характеристик для последующей смысловой интерпретации, анализа и обработки.

С точки зрения искусственного интеллекта, задача идентификации рисков может решаться как задача поиска аномалий в исторических массивах данных о деятельности, относящейся к области применения риск-менеджмента. Аномальные наблюдения в таких данных могут объясняться в том числе наличием взаимосвязей и взаимодействий между объектами и субъектами деятельности, уже приводящих к наступлению скрытых (еще не идентифицированных) рисков ситуаций и соответствующих последствий, либо являющихся потенциальными источниками возникновения таких ситуаций в будущем.

Подходы к риск-менеджменту, основанные на использовании методов поиска аномалий, активно развиваются и применяются, например, при выявлении мошеннических действий в онлайн-платежах [Domenig, 2021; Sekar, 2022], при управлении рисками инвестиционных портфелей [Westerlind, 2021] и при обнаружении сетевых атак [Sirag, Awadelkariem, 2022].

Авторами статьи разработан комплексный подход к автоматизированной идентификации новых типов рисков, основанный на методах поиска аномалий. Предлагаемый подход состоит из трех основных этапов:

- 1) выявление аномалий в историческом массиве данных;
- 2) разделение выявленных аномалий на однородные кластеры;
- 3) профилирование кластеров аномалий как потенциально новых типов рисков.

Для выделения аномалий (Anomaly Detection) могут использоваться различные методы интеллектуального анализа, в результате применения которых выявляются редкие объекты или события, значительно отличающиеся от большинства наблюдаемых объектов [Zimek, Schubert, 2017]. Обособленно данные методы имеют слабые систематические преимущества одного метода перед другими, также их эффективность сильно зависит от набора обучающих данных [Campos et al, 2016]. Следовательно, оптимальный подход к обнаружению аномалий должен сочетать в себе различные комбинации известных методов. В рамках рассматриваемого подхода предлагается использовать ансамблирование [Чесноков, 2018] методом голосования следующих наиболее эффективных методов поиска аномалий:

- 1) метода трех сигм [Pukelsheim, 1994];
- 2) метода эллипсоидальной аппроксимации [Rousseeuw, Driessen, 1999];
- 3) метода локального уровня выброса [Breunig, et al, 2000];
- 4) метода изолирующего леса [Liu, Ting & Zhou, 2008].

Для формирования однородных кластеров аномалий предлагается использовать кластерный анализ, который реализуется с помощью интеллектуальных алгоритмов машинного обучения без учителя: k-means, иерархический алгоритм кластеризации, Birch. Поскольку в задаче разделения аномалий на однородные группы нет априорно известного количества кластеров и распределения по кластерам, для выбора лучшего метода кластерного анализа и настройки его параметров могут использоваться внутренние метрики качества кластеризации: силуэт, индекс Дэвиса-Болдуина, индекс Калинского-Харабаша [Шалымов, 2008].

Профилирование кластеров аномалий заключается в описании характерных признаков, выделяющих данную группу аномалий от других групп. Для профилирования каждого аномального кластера выполняются следующие шаги:

- 1) выделение индикаторов кластера аномалий методами корреляционного анализа;
- 2) расчет показателей распределения количественных индикаторов кластера (средние, квантили, минимальные и максимальные значения);
- 3) расчет частот распределения категориальных индикаторов кластера.

Разработанный профиль аномального кластера представляет собой сущность, описывающую потенциально новый тип риска и предназначенную для последующих процедур смысловой интерпретации, валидации на новых данных, включения в типологию риск-менеджмента предметной области и встраивания в принятые организацией процессы оценки рисков, в том числе с использованием изложенных далее подходов.

##### ***Оценка вероятности наступления рисков***

Задача оценки вероятности наступления риска может рассматриваться в качестве частного случая классификации и решаться с помощью моделей искусственного интеллекта, позволяющих определять вероятность принадлежности объекта к тому или иному классу. Классификация относится к задачам обучения с учителем, где обучающая выборка состоит из множества  $X$ , содержащего признаки объектов, и соответствующего множества  $Y$ , содержащего метки классов, к которым относятся данные объекты. В

общем виде задача классификации является построением алгоритма отображения множества  $X$  в множество  $Y$ , способного сопоставить произвольный объект с соответствующими ему метками классов.

Таким образом, задача оценки вероятности наступления риска сводится к задаче бинарной классификации. В качестве выборки данных для обучения классификационной модели используются накопленные исторические массивы информации о зафиксированных фактах наступления рисков в прошлом, а также о характеристиках субъектов и объектов, имеющих отношение к наступившим рискам. Обученная на таких данных классификационная модель определяет вероятность наступления риска определенного типа в будущем по совокупности характеристик субъектов и объектов, относящихся к оцениваемой ситуации. При этом вероятность наступления риска фактически является вероятностью принадлежности объекта к классу  $A$ , где  $A$  – класс наличия риска, а  $B$  – класс отсутствия риска.

Если необходимо определять вероятность наступления нескольких типов рисков, возможно применение двух способов:

1) применение модели мультиклассовой классификации (когда каждый объект из множества  $X$  может относиться одновременно к нескольким классам);

2) применение нескольких бинарных классификаторов, каждый из которых максимально адаптирован для наиболее точной оценки одного конкретного типа риска.

В случаях, когда риски разнородны по своей природе и зависят от разного набора признаков, наиболее эффективным видится способ обучения нескольких бинарных классификаторов (для каждого типа риска) с последующим вычислением интегральной оценки вероятности наступления риска.

При оценке вероятности наступления рисков с использованием классификаторов следует учитывать ряд ключевых особенностей, характерных для данного класса задач. Первая особенность заключается в том, что для обучающих выборок, используемых в задачах прогнозирования вероятности риска, характерна крайняя несбалансированность, то есть соотношение объектов, принадлежащих к классу  $B$  (отсутствие риска), к объектам класса  $A$  (наличие риска) обычно находится в диапазоне между 100:1 и 10000:1 [Leevy, et al, 2018]. Данный фактор указывает на необходимость уделить особое внимание балансировке обучающей выборки и выбору специфических (наиболее релевантных) метрик качества обучения модели.

Для балансировки обучающей выборки может быть использован метод «under-sampling», основывающийся на выборочном исключении из обучающего набора данных объектов преобладающего класса, или метод «over-sampling», базирующийся на генерации синтетических объектов минорных классов и их добавлении в обучающий набор данных.

Как было отмечено выше, при обучении классификаторов на несбалансированных выборках необходимо использовать релевантные метрики качества формируемых моделей. Например, такая метрика оценки как «точность» («accuracy»), может оценивать качество модели более чем в 99% даже в тех случаях, когда формируемые классификатором прогнозы о принадлежности объектов к минорному классу в большинстве случаев являются ошибочными. Поэтому корректнее использовать более информативные метрики оценки качества классификаторов, например, такие как метрика «сбалансированная точность» («balanced accuracy»), оценивающая долю корректных ответов модели с учетом дисбаланса классов в выборке, метрика точности «precision», метрика полноты («recall») и метрика «F-мера» – среднее гармоническое значение точности и полноты. Кроме того, отслеживание баланса долей ложноположительных и ложноотрицательных предсказаний позволяет настроить модель под специфику конкретной прикладной задачи оценки вероятности наступления рисков. Так, например, в некоторых задачах большое количество ложноположительных срабатываний может нести в себе определенные финансовые издержки, при том, что редкая ложноотрицательная оценка риска является не столь критичной. И напротив, в таких задачах как, например, оценка риска в медицинской диагностике, модели настраиваются в соответствии с максимальной осторожностью для исключения (минимизации) потенциальных пропусков опасных (несущих риск) объектов.

Еще одной особенностью, которую необходимо учитывать, является возможность случайных и преднамеренных, обусловленных человеческим фактором ошибок при фиксации в историческом массиве данных фактов наступления рисков. Наибольшую проблему здесь представляют ситуации, когда в целях преднамеренного сокрытия факта наступления риска (нарушения) объект отмечается человеком (оператором) как объект без риска. Это приводит к ситуации наличия скрытых не идентифицированных рисков. В историческом массиве накапливается искаженная недостоверная информация, что существенно снижает качество моделей оценки вероятности наступления рисков, обучаемых на данном массиве. Одной из эффективных мер по нивелированию данного фактора является рассмотренный ранее в данной статье подход к автоматизированной идентификации новых типов рисков через выявление аномалий. Аномальные объекты могут исключаться из обучающей выборки для классификаторов по известным (идентифицированным) типам рисков либо для однородных групп (кластеров) выявленных аномалий как потенциально новых типов рисков могут обучаться отдельные классификаторы (во втором варианте результаты идентификации новых типов рисков фактически являются дополнительной разметкой набора данных для обучения классификационных моделей).

Для оценки вероятности наступления риска могут использоваться как «классические методы» машинного обучения, так и глубокие нейронные сети. Оптимальная архитектура нейронной сети в данном случае подбирается в соответствии с природой данных из обучающей выборки: для оценки вероятности

наступления риска могут использоваться как полносвязные нейронные сети, так и, к примеру, сверточные или рекуррентные.

Детальное рассмотрение архитектур нейронных сетей и моделей машинного обучения, пригодных для решения подобных задач, выходит за рамки данной статьи. Перечислим лишь наиболее распространенные и применимые модели классификаторов [Sarker, 2021], позволяющих определить не только принадлежность оцениваемого объекта к классу риска, но и вероятность принадлежности к этому классу:

- 1) логистическая регрессия;
- 2) ближайшие соседи;
- 3) решающие деревья;
- 4) случайный лес;
- 5) градиентный бустинг.

Практика показывает, что эффективность перечисленных моделей классификации зависит от специфики конкретной решаемой задачи и природы данных. Поэтому в большинстве случаев целесообразно обучать несколько различных моделей с различными конфигурациями, а выбор наиболее оптимальной модели и ее гиперпараметров производить по метрике качества, полученной на тестовой выборке данных.

На сегодняшний день применение моделей искусственного интеллекта для оценки вероятности наступления рисков наиболее распространено в следующих сферах:

1. **Финансы.** Модели искусственного интеллекта успешно применяются для решения задач кредитного скоринга [Chow, Jacky, 2017]. Также интеллектуальные модели успешно применяются для анализа финансовых транзакций на предмет мошенничества и для оценки страховых рисков [Tiwari, 2021; Eling, Nuessle, & Staubli, 2021].

2. **Здравоохранение.** Модели искусственного интеллекта позволяют прогнозировать вероятность риска возникновения и обострения заболеваний, а также сопутствующих им мер, таких как незапланированная повторная госпитализация, приводящая к снижению уровня ухода за пациентами и перегрузке системы здравоохранения [Somya, et al, 2022].

3. **Государственное управление.** В качестве примера применения методов искусственного интеллекта для управления рисками в сфере государственного управления можно привести модель автоматизированной оценки риска государственных контрактов, направленную на повышение эффективности государственных закупок и оптимизацию бюджетных расходов [Елисеев., Романов, 2018]. **Ошибка! Закладка не определена.**

4. **Образование.** В образовательной сфере применение методов искусственного интеллекта также может быть достаточно разнообразным и применяться, например, для анализа вероятности риска отчисления студента [Medina, 2021] или неудовлетворительных результатов академической успеваемости [Lakkaraju, et al, 2015].

5. **Безопасность.** Интеллектуальные модели могут использоваться для выявления административных и уголовных правонарушений и их своевременного предотвращения [Fonseca, Pinto, Sargento, 2021].

Резюмируя вышеизложенное, можно сказать, что методы искусственного интеллекта применимы в задачах оценки вероятности наступления рисков и имеют потенциал более эффективного и качественного (по сравнению с классическими технологиями<sup>14,17</sup> решения этих задач за счет снижения нагрузки на экспертов и снижения влияния человеческого фактора на процесс и результат оценки рисков.

#### **Оценка величины вероятных последствий при наступлении риска**

Величина вероятных последствий при наступлении риска как правило может быть выражена количественными значениями, которые зависят от ряда факторов, характеризующих сам риск, а также состояние взаимосвязанных с ним процессов, объектов и субъектов в момент возникновения риска. Поэтому задача оценки величины вероятных последствий при наступлении риска может рассматриваться в качестве частного случая задачи регрессии и решаться с использованием соответствующих моделей искусственного интеллекта. Регрессия, как и классификация, относится к задачам обучения с учителем, где обучающая выборка состоит из множества  $X$ , содержащего признаки объектов, и соответствующего множества  $Y$ , содержащего вещественные числа. В общем виде задача регрессии является построением алгоритма отображения множества  $X$  на множество  $Y$ , способного сопоставить произвольный объект с вещественным числом.

Применительно к рассматриваемой задаче во множестве  $X$  обучающей выборки для построения регрессионной модели содержатся факты наступления рисков в прошлом, а также значения признаков, характеризующих случившиеся риски. Множество  $Y$  содержит значения величин, количественно характеризующих последствия наступившего риска (например, сумму убытков и издержек организации, сумму начисленных штрафов).

Также задача регрессии может быть рассмотрена как задача количественного расчета риска в менее конкретном смысле, когда необходимо не просто рассчитать вероятные материальные издержки, а провести прогнозирование значения, отражающего уровень какой-либо величины или степень риска. В данном случае зачастую требуется проводить дополнительную интерпретацию полученных выводов модели – сопоставлять прогнозные значения с некоторыми критериями риск-категорирования, принятыми

в области деятельности организации. Как правило, такие критерии формируются на основе специальных экспертных оценок и закрепляются в отраслевых документах (нормативных актах, стандартах, рекомендациях)<sup>18</sup>.

Регрессионные модели могут сочетаться с рассмотренными выше по тексту классификационными моделями оценки вероятности наступления рисков. Система искусственного интеллекта, использующая и классификацию, и регрессию, может формировать интегральную оценку риска, состоящую из вероятности его наступления и количественной оценки величины возможных последствий.

Аналогично оценке вероятности наступления риска с применением классификаторов, для оценки вероятных последствий при наступлении риска могут использоваться как «классические методы» машинного обучения (например, линейная регрессия, метод ближайших соседей, решающие деревья, случайный лес, градиентный бустинг), так и глубокие нейронные сети.

Рассмотренный подход уже находит эффективное применение, например, в медицине и финансах. В области медицины регрессионные модели могут использоваться для прогнозирования значений медицинских показателей, изменением которых сопровождается возникновение того или иного заболевания [Aishah, et al, 2018]. В финансовой сфере регрессионные методы могут быть использованы при построении систем, прогнозирующих ущерб в материальном эквиваленте для задач микроэкономического и макроэкономического моделирования [Dzedik &, Lontsih, 2018].

#### **Применение интеллектуальных систем на основе онтологии предметной области**

Стандартные процессы риск-менеджмента<sup>19,20</sup> требуют принятия своевременных решений, основанных на полных, достоверных и актуальных данных из различных источников, в том числе из внутренних информационных систем организации, информационных систем партнеров, государственных информационных систем. Эти источники разнородны. Их состав и структура передаваемой информации подтверждены динамичным изменением под влиянием различных факторов, включая факторы развития организации. На эти же источники и данные опираются рассмотренные выше интеллектуальные модели оценки рисков.

Поэтому крайне важными в области риск-менеджмента являются задачи интеграции разнородной информации, ее эффективной систематизации и формализации.

В качестве методологического фундамента решения этих задач возможно использование онтологии – формализованного описания предметной области деятельности, входящей в область применения риск-менеджмента. На текущий момент видится несколько наиболее перспективных направлений применения средств онтологии:

1) формирование модели предметной области, определяющей терминологический аппарат, логические и физические модели данных, а также поддерживающей надежные механизмы актуализации и валидации с использованием средств онтологии;

2) создание и ведение формализованных правил применения интеллектуальных моделей оценки различных типов рисков к различным объектам и субъектам деятельности в зависимости от их характеристик;

3) создание и ведение формализованных правил обучения моделей оценки конкретных типов рисков, включая правила отбора данных из различных, определенных в физической модели, источников для формирования обучающих выборок, правила предварительной обработки этих данных, перечни применимых математических алгоритмов.

#### **Выводы**

В настоящей статье проведен обзор ключевых задач в области управления рисками, предложены и рассмотрены подходы к их решению с использованием современных технологий искусственного интеллекта и анализа больших данных.

Для идентификации и выявления новых типов рисков разработан автоматизированный подход, основанный на поиске в историческом массиве данных аномалий с использованием техники голосования ансамблем методов и последующем разделении выявленных аномалий на однородные кластеры, соответствующие потенциально новым типам рисков. Данный подход является новым и может применяться в риск-менеджменте различных сфер деятельности.

Оценка вероятности наступления известных типов рисков сводится к задаче классификации и, соответственно, может решаться интеллектуальными классификационными моделями с применением классических методов машинного обучения или нейронных сетей. Залогом эффективного применения данных моделей являются:

1) предварительная балансировка обучающего набора данных с применением технологий «under-sampling» или «over-sampling»;

<sup>18</sup> ГОСТ 12.1.007–76. Межгосударственный стандарт. Система стандартов безопасности труда. Вредные вещества. Классификация и общие требования безопасности

<sup>19</sup> ISO 31000:2018 «Risk management – Guidelines»

<sup>20</sup> ГОСТ Р ИСО 31000–2019 «Национальный стандарт Российской Федерации. Менеджмент риска. Принципы и руководство»

2) выбор объективной метрики качества модели, характеризующей допустимый уровень ложноположительных или ложноотрицательных срабатываний с учетом специфики решаемой прикладной задачи;

3) нивелирование ошибок разметки обучающего набора данных, обусловленных человеческим фактором, за счет применения предложенного в статье подхода к автоматизированной идентификации новых типов рисков через выявление аномалий.

Количественная оценка величины вероятных последствий наступления риска может осуществляться интеллектуальными регрессионными моделями. Совместное применение интеллектуальных классификационных и регрессионных моделей позволяет анализировать риск комплексно и формировать его интегральную оценку, включающую вероятность наступления риска и оценку величины последствий.

Интеллектуальные системы, построенные на основе онтологии предметной области, позволяют повысить качество формализации информации и знаний организации, необходимых для процессов управления рисками, а также повысить эффективность процессов построения и сопровождения интеллектуальных моделей оценки рисков.

Обозначенные авторами подходы являются универсальными и встраиваемыми в стандартные процессы риск-менеджмента организации, а также имеют существенный потенциал применения для оптимизации механизмов управления рисками, заключающийся в снижении нагрузки на экспертов, повышении надежности, эффективности и результативности соответствующих процессов.

Внедрение рассмотренных подходов не только своевременно, но становится и ключевым условием развития и конкурентоспособности в различных отраслях экономики и сферах государственного регулирования.

#### Литература

1. Елисеев Д. А., Романов Д. А. Машинное обучение: прогнозирование рисков госзакупок. Открытые системы. СУБД. 2018
2. Чесноков М.Ю. Поиск аномалий во временных рядах на основе ансамблей алгоритмов DBSCAN // Искусственный интеллект и принятие решений. – 2018. – № 1. – С. 99—107.
3. Шалымов Д.С. Алгоритмы устойчивой кластеризации на основе индексных функций и функций устойчивости // Стохастическая оптимизация в информатике. СПб.: Изд-во С.-Петербургского университета. 2008. 4. С. 236-248.
4. A F Q A Aishah, K Ummu J, M R Zainuriah and A K Norhilda - Multiple linear regression model analysis in predicting fasting blood glucose level in healthy subjects IOP Conference Series: Materials Science and Engineering, Volume 469, 1st International Postgraduate Conference on Mechanical Engineering (IPCME2018) 31 October 2018, UMP Pekan, Pahang, Malaysia.
5. Breunig M. M., Kriegel H.-P., Ng R. T., Sander J. LOF: Identifying Density-based Local Outliers. Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data. 2000. pp. 93–104. doi:10.1145/335191.335388.
6. Campos G.O., Zimek A., Sander J., Campello R., Micenkova B., Schubert E., Assent I., Houle M.E. On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study // Data Mining and Knowledge Discovery. 2016. v. 30, No. 4. doi:10.1007/s10618-015-0444-8.
7. Chow, Jacky. (2017). Analysis of Financial Credit Risk Using Machine Learning. 10.13140/RG.2.2.30242.53449.
8. Dzedik V.A., Lontsikh P.A. Using machine-learning methods to analyze economic loss function of quality management processes (2018) J. Phys.: Conf. Ser. 1015 032031.
9. DomenigT, Zvizdic E, Vanini P (2021). Online Payment Fraud: From Anomaly Detection to Risk Management. Operational Risk and Cyber Risk
10. Eling, M., Nuessle, D. & Staubli, J. (2021) The impact of artificial intelligence along the insurance value chain and on the insurability of risks. Geneva Pap Risk Insur Issues Pract (2021). <https://doi.org/10.1057/s41288-020-00201-7>
11. Leevy, J.L., Khoshgoftaar, T.M., Bauder, R.A. et al. A survey on addressing high-class imbalance in big data. J Big Data 5, 42 (2018). <https://doi.org/10.1186/s40537-018-0151-6>
12. Liu F.T., Ting K.M., Zhou Z.-H. Isolation forest. In Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on (pp. 413-422). doi: 10.1109/ICDM.2008.17.
13. Lakkaraju H, Aguiar E, Shan C, Miller D, Bhanpuri N, Ghani R, and Addison K.L. (2015). A Machine Learning Framework to Identify Students at Risk of Adverse Academic Outcomes. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15). Association for Computing Machinery, New York, NY, USA, 1909–1918. DOI: <https://doi.org/10.1145/2783258.2788620>
14. Medina B, Mauricio & Meza Narváez, Claudia & Gálvez-García, Germán. (2021). Prediction of student attrition risk using machine learning - Journal of Applied Research in Higher Education. 10.1108/JARHE-02-2021-0073
15. Mohanty S. D., Lekan D, McCoy T.P., Jenkins M, Prashanti Manda Machine learning for predicting readmission risk among the frail: Explainable AI for healthcare, Patterns, Volume 3, Issue 1 (2022) <https://doi.org/10.1016/j.patter.2021.100395>
16. Pukelsheim F. The Three Sigma Rule. The American Statistician, 48(2), 1994. pp. 88–91. <https://doi.org/10.2307/2684253>

17. Rousseeuw P.J., Driessen V.K.: A fast algorithm for the minimum covariance determinant estimator. *Technometrics*. 1999. Vol. 41 (3). pp. 212-223.
18. Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN COMPUT. SCI.* 2, 160 (2021). <https://doi.org/10.1007/s42979-021-00592-x>
19. Sekar M. (2022) Fraud and Anomaly Detection. In: *Machine Learning for Auditors*. Apress, Berkeley, CA. [https://doi.org/10.1007/978-1-4842-8051-5\\_21](https://doi.org/10.1007/978-1-4842-8051-5_21)
20. Sirag H., Awadelkariem S.D. (2022) A Review on Intrusion Detection System Using a Machine Learning Algorithms. *Proceedings of International Conference on Emerging Technologies and Intelligent Systems. ICETIS 2021. Lecture Notes in Networks and Systems*, vol 322. Springer, Cham. [https://doi.org/10.1007/978-3-030-85990-9\\_24](https://doi.org/10.1007/978-3-030-85990-9_24)
21. Tiwari, P., Mehta, S., Sakhuja, N., Kumar, J., & Singh, A. K. (2021). Credit Card Fraud Detection using Machine Learning: A Study. <https://doi.org/10.48550/arXiv.2108.10005>
22. Zimek A., Schubert E. *Outlier Detection // Encyclopedia of Database Systems*. – Springer New York, 2017. DOI: 10.1007/978-1-4899-7993-3\_80719-1.
23. Westerlind S (2018) *Anomaly Detection for Portfolio Risk Management*. KTH Industrial Engineering and Management Industrial Management. Stockholm

#### References in Cyrillics

1. Eliseev D. A., Romanov D. A. Mashinnoe obuchenie: prognozirovaniye riskov goszakupok. *Ot-kry`ty`e sistemy` . SUBD*. 2018
2. Chesnokov M.Yu. Poisk anomalij vo vremenny`x ryadax na osnove ansamblej algoritmov DBSCAN // *Iskusstvenny`j intellekt i prinyatie reshenij*. – 2018. – № 1. – S. 99—107.
3. Shaly`mov D.S. Algoritmy` ustojchivoj klasterizacii na osnove indeksny`x funkcij i funk-cij ustojchivosti // *Stoxasticheskaya optimizaciya v informatike*. SPb.: Izd-vo S.-Peterburgskogo universiteta. 2008. 4. S. 236-248.

Гоглев Н.Н., руководитель направления, [Nikita.Goglev@centre-it.com](mailto:Nikita.Goglev@centre-it.com);  
 Е.В. Касаткина, к.ф.-м.н., доцент, аналитик, [Ekaterina.Kasatkina@centre-it.com](mailto:Ekaterina.Kasatkina@centre-it.com) .  
 С.А. Мигалин, ведущий системный аналитик, [Sergey.Migalin@centre-it.com](mailto:Sergey.Migalin@centre-it.com)  
 О.И. Муштак, аналитик, [Oksana.Mushtak@centre-it.com](mailto:Oksana.Mushtak@centre-it.com)  
 Ижевский государственный технический университет имени М.Т. Калашникова,  
 ул. Студенческая, 7, Ижевск, 426069, РФ

**Ключевые слова** – риск-менеджмент, управление рисками, идентификация рисков, оценка рисков, искусственный интеллект, машинное обучение, поиск аномалий, классификация, регрессия, онтология предметной области.

***Nikita Goglev, Ekaterina Kasatkina, Sergey Migalin, Oksana Mushtak, Risk management using modern artificial intelligence technologies and big data analysis***

#### Keywords

risk management, risk management, risk identification, risk assessment, artificial intelligence, machine learning, anomaly search, classification, regression, domain ontology

DOI: 10.34706/DE-2022-02-05

JEL classification C53 – Методы прогнозирования • Методы моделирования,

#### Abstract

The high level of uncertainty in the modern world, combined with the active development of artificial intelligence technologies and big data analysis, indicate the relevance and timeliness of the use of these technologies to improve the efficiency of processes in such an aspect of public and corporate governance as risk management. The introduction of risk management approaches based on the use of modern technologies is actually becoming one of the key conditions for the development of competitiveness in various sectors of the economy and spheres of state regulation. This article discusses the processes and tasks of risk management established by international and Russian risk management standards. The approaches that allow, in contrast to classical technologies, to more effectively solve the basic standard tasks of risk management through the use of modern technologies of artificial intelligence and big data analysis are formulated. Since methods for identifying new types of risks are poorly developed at the moment, the authors of the article have developed a new integrated approach to automated identification of new types of risks based on the use of intelligent methods for searching for anomalies in historical data arrays about the organization's activities. An approach to assessing the probability of occurrence of known risks and the magnitude of the likely consequences of their occurrence, based on the use of intelligent classification and regression models, is considered. Possible options for the practical application of intellectual systems based on the ontology of the subject area are formulated to improve the efficiency and quality of systematization of heterogeneous information related to risk management processes, including through the combination of ontological tools with models of artificial intelligence.