

УДК 004.89
DOI 10.52575/2687-0932-2022-49-4-820-832

Использование методов машинного обучения в задачах принятия решений при обеспечении качества в приборостроении

Конев К.А.

Уфимский университет науки и технологий,
Россия, 450076, Республика Башкортостан, г. Уфа, ул. Заки Валиди, д. 32
E-mail: sireo@rambler.ru

Аннотация. Обеспечение качества – важная и многогранная деятельность, оказывающая значительное влияние на конкурентоспособность и продукцию организации, на её системы управления в целом. Данная деятельность не является ценным и основным процессом, поэтому управляется как центр затрат, что создаёт объективные сложности в повышении её эффективности. Несмотря на большое число публикаций в сфере обеспечения качества, тематика поддержки принятия решений в данной области исследована слабо и рассматривается главным образом для руководителей. Концептуально проблема повышения эффективности решений рассмотрена в ряде публикаций автора. В работе данный подход рассматривается для специалистов по обеспечению качества уровня исполнителя или эксперта, и в такой постановке задачи используется впервые. В соответствии с ранее предложенной методологией целью исследования является формирование ситуационно-онтологической модели конкретной практической задачи выявления причины дефекта для приборостроительного предприятия методом машинного обучения (k-ближайших соседей). Результаты исследования позволили сделать вывод о возможности применения ситуационно-онтологической методологии с целью создания системы поддержки принятия решений для решения задач в типовой ситуации поиска причины дефекта на основе метода k-ближайших соседей для инженера по качеству приборостроительного предприятия.

Ключевые слова: поддержка принятия решений; машинное обучение; цифровой след; обеспечение качества; диаграмма Исикавы; классификация причин дефектов; авиационное приборостроение; метод k-ближайших соседей; Knime Analytics Platform

Для цитирования: Конев К.А. 2022. Использование методов машинного обучения в задачах принятия решений при обеспечении качества в приборостроении. Экономика. Информатика, 49(4): 820–832. DOI 10.52575/2687-0932-2022-49-4-820-832

Use of Machine Learning Methods in Decision Making to Ensure Quality in Instrument Manufacturing

Konstantin A. Konev

Ufa University of Science and Technology,
32, Zaki Validi St., Ufa, Republic of Bashkortostan, 450076, Russia
E-mail: sireo@rambler.ru

Abstract. Quality assurance is an important and various activity that has a significant impact on the competitiveness of both the organization's products and its management system as a whole. This activity is not a value-creating main process and is managed as a cost center, which creates objective difficulties in increasing its efficiency. Despite the large number of publications in the field of quality assurance, the topic of decision support in this area has been poorly researched and is considered mainly for high-level managers. Conceptually, the problem of increasing the efficiency of decisions is considered in a number of

publications of the author. In this paper, this approach is considered for quality assurance specialists at the level of the executer or expert and is used for the first time in this formulation of the problem.

The aim of the study is, in accordance with the previously proposed methodology, the formation of a situational-ontological model for solving a specific practical problem of identifying the cause of a defect for an instrument-making enterprise using the machine learning method (k-nearest neighbors).

The results of the study allowed us to conclude that it is possible to use the situational ontological methodology to create a decision support system for solving problems in a typical situation of finding the cause of a defect based on the k-nearest neighbors method for a quality engineer of an instrument-making enterprise.

Keywords: decision support; machine learning; digital footprint; quality assurance; Ishikawa diagram; classification of causes of defects; aviation instrumentation; k-nearest neighbor method; Knime Analytics Platform.

For citation: Konev K.A. 2022. Use of Machine Learning Methods in Decision Making to Ensure Quality in Instrument Manufacturing. Economics. Information technologies, 49(4): 820–832. DOI 10.52575/2687-0932-2022-49-4-820-832

Введение

Обеспечение качества – многогранная и разнообразная деятельность, во многом определяемая спецификой бизнеса. Актуальность исследований в данной области продиктована сложившимися обстоятельствами, связанными с необходимостью возрождения целых отраслей по выпуску конкурентоспособной продукции. Важной особенностью данной деятельности является большое количество рисков и значительный объём принимаемых решений, что требует высочайшей компетентности и осведомлённости должностных лиц. При этом нужно учитывать, что обеспечение качества относится не к основной, а к вспомогательной деятельности и, соответственно, управляется как центр затрат. Возникает противоречие: требования к специалистам данного профиля высоки, а затраты на них ограничены. В качестве решения данной проблемы перспективным выглядит повышение эффективности решаемых задач без расширения штата сотрудников, то есть за счёт интенсификации их труда. Поскольку специфика задач таких специалистов связана с принятием решений, то подход к интеллектуализации их труда и проектированию систем поддержки принятия решений выглядит оправданным.

Важно отметить, что несмотря на большое число публикаций и существенные положительные сдвиги в практической области в сфере обеспечения качества, особенно в части повышения качества организационных взаимодействий (менеджмент качества), тематика поддержки принятия решений в данной области исследована слабо. Приоритет отдаётся задачам руководителя уровня директора по качеству, а не специалиста по качеству, контролёра, аудитора и т.д. Между тем, например, ошибка контролёра может в конечном итоге привести к авиационному происшествию или даже катастрофе, а аудитор может не заметить организационно-техническую проблему, приносящую миллионные убытки.

Кроме того, спектр задач обеспечения качества очень широк и требует привлечения специалистов различного профиля. Например, специалисты по техническому контролю осуществляют технический контроль деталей, агрегатов и готовых изделий, технологи обеспечивают надзор за технологичностью и культурой производства, метрологи обеспечивают точность измерительного инструмента и т.д., а есть ещё и ответственные за обеспечение качества в цехах, специалисты по учёту затрат, руководители всех уровней, которые также участвуют в данной деятельности. Таким образом, невозможно найти специалиста по всем аспектам обеспечения качества, но можно и нужно увязать множество функций и знаний различных специалистов и руководителей предприятия в единую систему, собрав и обобщив их опыт.

Способ, с помощью которого можно решить данную актуальную задачу, является использование современных технологий цифровой трансформации организации, таких

как машинное обучение и цифровой след для сбора и анализа опыта в предметной области обеспечения качества. Целью статьи является формирование концептуальной модели машинного обучения для предметной области обеспечения качества и структурной модели одной из её типовых задач на основе ситуационно-онтологической методологии [Антонов и др. 2021].

Ситуационно-онтологический подход

Концептуально ситуационно-онтологический подход основывается на методологии Захмана [Zachman, 1987]. Суть подхода в представлении бизнес-процесса, в рамках которого принимается решение, в виде иерархии моделей от верхнего уровня – уровня модели организации до уровня модели оценки принятого решения. Базовой моделью в данной концепции является дискретно-событийная ситуационная модель, которая описывает ситуацию принятия решения. При этом модели связаны как множества отношениями композиции, поэтому общая модель образует таксономию – разновидность онтологических моделей, что и предопределило название подхода.

Важным элементом концепции является инвариантность по методу принятия решений, т.е. он выбирается под уровень задач лица, принимающего решения в бизнес-процессе, а не предопределяется исходя из предпочтений исследователя, что позволяет повысить гибкость подхода. Вторым важным аспектом является упор на использование интеллектуальных методов и обязательное сохранение опыта в базе знаний.

Для решения задач в типовых ситуациях, в которых имеется возможность накапливать статистику, например, за счёт учёта цифрового следа, и имеются специалисты технического профиля достаточно высокой квалификации, хорошо подходят методы машинного обучения.

Машинное обучение и цифровой след в задачах принятия решений

Машинное обучение (англ. machine learning, ML) – класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение за счёт применения решений множества сходных задач [Флах, 2015]. Машинное обучение используется для решения трёх основных задач: классификации, прогнозирования и кластеризации. Классификация подразумевает задачу идентификации, т.е. отнесения объекта к заранее известному классу. Прогнозирование осуществляется за счёт регрессионного анализа и приближения оцениваемого параметра к некоторому значению, т.е. выполняется предсказание значения некоторой характеристики объекта, процесса или явления. Наконец, кластеризация направлена на выявление новых категорий (кластеров), к которым можно относить объекты.

Если рассматривать ситуацию, связанную с принятием решения, то машинное обучение может быть использовано в целом ряде таких задач, но для его применения необходимо использовать ещё одну современную информационную технологию, называемую цифровым следом (цифровой тенью). Цифровой след – это совокупность состояний и характеристик в этих состояниях, которые оставляет объект в каких-либо средах: документах, социальных сетях, чатах, научных конференциях и т.д. [Kim и др., 2018]. Данное понятие справедливо для любого сложного объекта или бизнес-процесса. Сложное изделие с большим жизненным циклом, такое как капитальное здание, автомобиль или самолёт также имеет множество разнообразных состояний и проявляют разные характеристики в них, только они наблюдаются и фиксируются иначе. На большой выборке они могут дать много полезной информации. По сути, речь идёт о формировании истории, онтологии состояний конкретного объекта и исследовании закономерностей изменения этих состояний на большой выборке. Если цифровому следу поставить в соответствие множество данных, достаточно полно и правильно отражающих соответствие физическому объекту (или процессу), то это будет уже синхронизированная с ним модель, т.е. цифровой двойник (digital

twin) [Negri, 2017], основным свойством которого будет возможность достаточно точного представления динамики жизненного цикла физического процесса, что позволит уже говорить и о «Больших данных» (Big Data), и об «интернете вещей» (IoT) во взаимосвязи, т.е. переходить к требованиям концепции «Индустрии 4.0» [Chen, 2018].

Теоретически собрать большой объём данных об объекте (процессе) – цифровой след при принятии решений никаких проблем нет, но на практике подобный подход требует большого объёма знаний и сложный механизм обработки данных о предметной области. Для решения задач принятия решений принято использовать интеллектуальные методы в рамках систем поддержки принятия решений [Антонов, Конев, 2021]. Укрупнённая структурная схема интеллектуальной подсистемы системы поддержки принятия решений показана на рисунке 1.

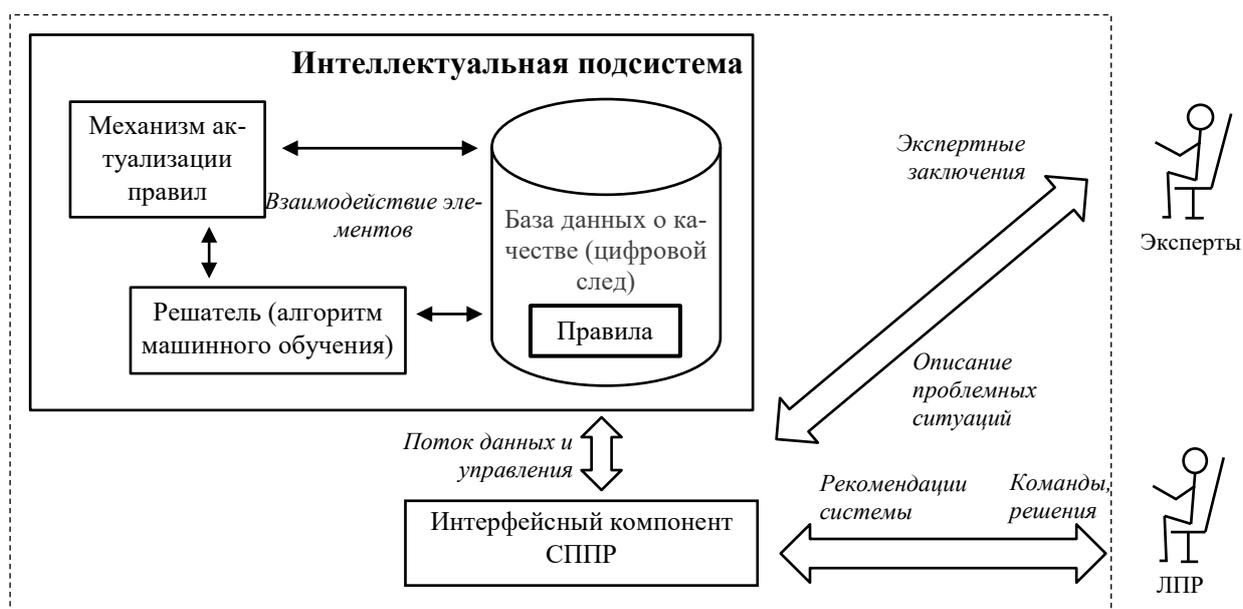


Рис. 1. Структура интеллектуальной подсистемы системы поддержки принятия решений при обеспечении качества

Fig. 1. The structure of the intellectual subsystem of the decision support system for quality assurance

Рассмотрим такой подход для сферы обеспечения качества на примере одной из типичных задач.

Задача выявления причины дефекта из перечня заранее известных

Согласно обзору заявлений предприятий от Дерекса Сноу [Snow, 2019] на 2019 год в сфере применения в промышленности метод машинного обучения имеет перспективы в четырёх сегментах: общее управление производством, техническое обслуживание, прогнозирование ошибок и обеспечение качества. Очевидно, что последние две сферы взаимосвязаны. Применительно к прогнозированию отказов оборудования и обнаружения аномалий в производстве исследователем отмечены области деятельности в производстве электронных компонентов, лекарственных средств и автомобилей. В целом примеров успешного применения методов машинного обучения в отрасли авиационного приборостроения ни у Сноу, ни у других исследователей обнаружить не удалось, хотя многие исследователи отмечают большие перспективы методов искусственного интеллекта в авиационной отрасли [Коробеев, Чучаев, 2018]. В данной связи проектирование системы поддержки принятия решений на основе машинного обучения для отрасли авиационного приборостроения является новой задачей.

Рассмотрим задачу принятия решения о причине дефекта на предприятии авиационного приборостроения как любой факт несоответствия установленным и предполагаемым требованиям, выявленный на любом этапе жизненного цикла продукции. Допустим, имеется множество типовых причин и множество признаков, по которым можно отнести конкретную ситуацию из множества наблюдаемых к конкретному типу причин, либо принять решение, что данный тип причин является новым. Визуализируем признаки этих ситуаций, для этого используем метод Исикавы [Исикава, 1988]. На рис. 2 показана типовая диаграмма Исикавы для производственного предприятия.

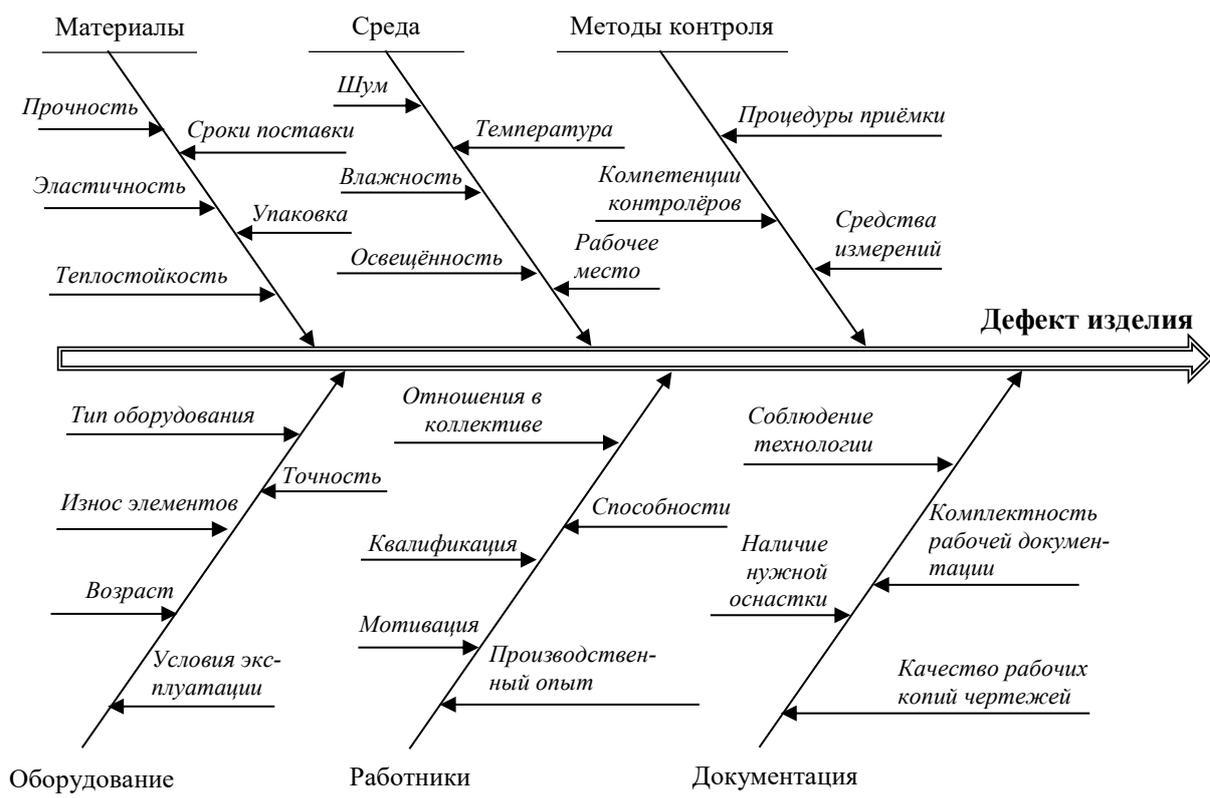


Рис. 2. Диаграмма Исикавы для определения причины дефекта
 Fig. 2. Ishikawa diagram to determine the cause of a defect

Каждая из показанных на рис. 2 характеристик должна быть оцифрована, т.е. получить некоторую шкалу оценки, которую должен использовать эксперт. На основе такой оценки какие-то признаки будут полностью отброшены, а какие-то выбраны в виде перечня кандидатов в причины. Принятие решений в данной ситуации имеет 2 этапа: классификации и выбора. На этапе классификации осуществляется поиск той ветки в диаграмме Исикавы, которой лучше всего соответствует наблюдаемый дефект. На этапе выбора осуществляется выбор решения по действиям с причиной данного дефекта.

Каждое обращение к поиску причины оставляет цифровой след, который можно запоминать и статистически обрабатывать. Если в течение заданного времени, например, трёх месяцев, для конкретного дефекта причина не будет исправлена, то выбор можно считать удачным и запоминать как успешный прецедент для модели поиска причины, иначе прецедент будет помечен как неудачный. После каждой ситуации выбора можно анализировать набор параметров, которые сформируют правила для обучения алгоритма выбора. Схема работы системы поддержки принятия решений о причинах дефектов показана на рисунке 3.

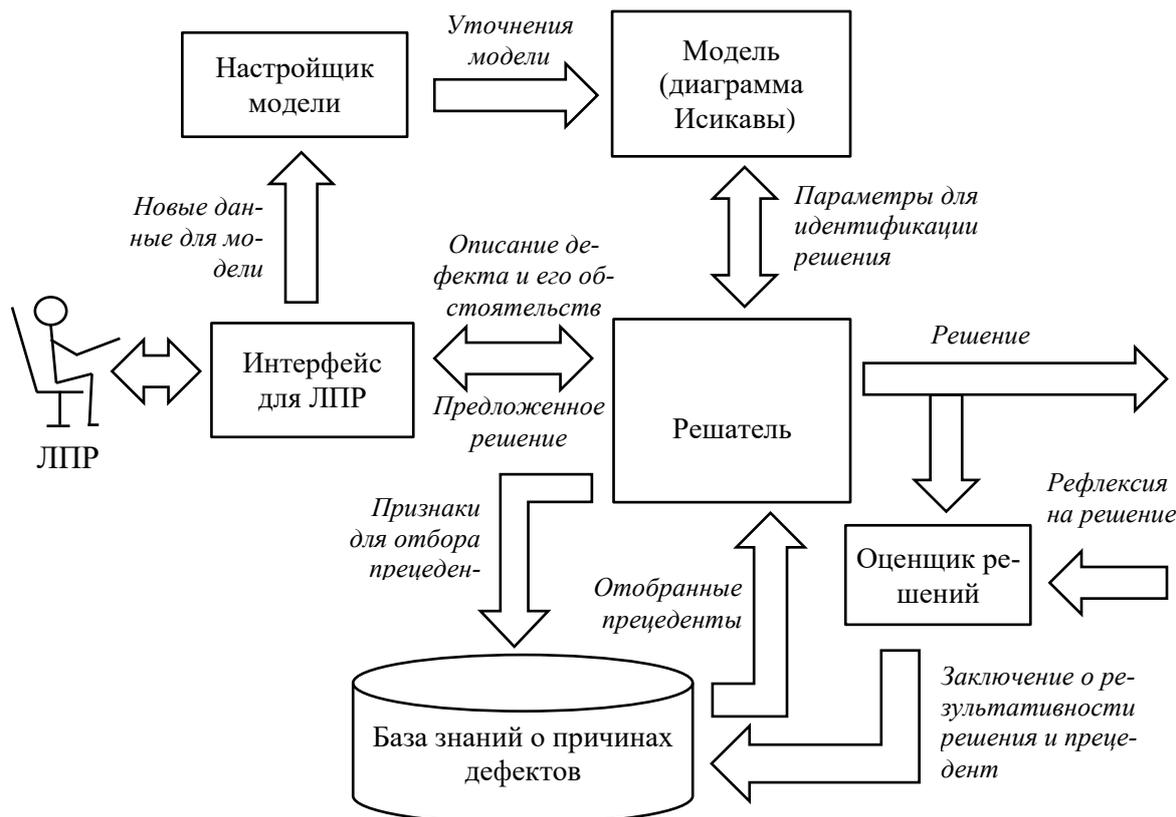


Рис. 3. Схема работы системы поддержки принятия решений о причинах дефектов

Fig. 3. Scheme of operation of the decision support system on the causes of defects

Фактически предлагается алгоритм выбора ветки диаграммы Исикавы на основе ранее сделанных выборов, которые оценены по наличию / отсутствию опровержений по выбранному варианту решения. Возникает вопрос о месте машинного обучения в данной схеме. Выбор ветки диаграммы Исикавы является задачей идентификации, т.е. типичной задачей алгоритмов машинного обучения, в которой необходимо статистически обработать набор лингвистических признаков, которые содержатся в описании дефекта.

Рассмотрим набор таких признаков для задачи выбора причины дефекта. В их число могут, например, входить:

- место обнаружения дефекта;
- стадия процесса, на которой выявлен дефект;
- операция технологического процесса, на которой сделана ошибка;
- ответственность за операцию, на которой сделана ошибка;
- комплектность технологической документации на рабочем месте;
- полнота описания конструкции в чертежах конструкторской документации;
- исправность оборудования на рабочем месте;
- исправность инструмента, приборов на рабочем месте;
- исправность оснастки на рабочем месте;
- сроки поверки мерительного инструмента на рабочем месте;
- стабильность цепей электропитания на рабочем месте;
- соответствие температуры, влажности и освещённости на рабочем месте;
- уровень статического напряжения на рабочем месте;
- уровень чистоты и культуры производства на рабочем месте;
- уровень эргономичности рабочего места;
- соответствие рабочей одежды требованиям рабочего места;

- исправность тары для хранения и межцеховой транспортировки деталей, сборочных единиц и готовой продукции;
- соответствие мест хранения деталей, сборочных единиц и готовой продукции;
- соответствие содержания выданных рабочему компонент для работы по составу и номиналу комплектовочной ведомости и требованиям конструкторской документации;
- соответствие лакокрасочных материалов требованиям конструкторской документации и т.д.

Для реального бизнес-процесса таких показателей может быть несколько десятков, причём они сильно увязаны с особенностями бизнес-процессов.

Обычно при расследовании причин дефектов ЛПР должен лично анализировать данные признаки. Между тем, если в организации внедрена цифровая система производственного учёта, то значительный объём таких данных можно собирать из этой системы практически автоматически.

Цифровой след при формировании информации для обработки методами машинного обучения

Выше затрагивался вопрос цифрового следа для сбора данных в промышленности при решении задач принятия решений. Применительно к задаче о классификации причин дефектов на приборостроительном предприятии, цифровой след формируется в нескольких базах данных и прочих источниках. К ним относятся данные об:

- оборудовании, его обслуживании и метрологическом обеспечении;
- работниках, их квалификации, взаимоотношениях и условиях труда;
- конструкторской и технологической документации, их соответствии требованиям, актуальности и качеству копий;
- комплектующих, материалах и их состоянии;
- методах контроля;
- проведённых контрольных мероприятиях и выявленных дефектах.

Рассмотрим схему сбора данных для цифрового следа для рассмотренной выше задачи. В общих чертах она включает следующие действия:

- идентификация объекта исследования, которым в нашем случае является изделие;
- выявление состояний, в которых может находиться изделие (соответствие или дефект);
- описание характеристик, которые изделие может демонстрировать в каждом состоянии;
- формирование метамоделей, т.е. правил свертки данных о характеристиках в состояниях объекта и критериев для интерпретации полученных значений (методы идентификации причин дефектов);
- сбор информации о состояниях и характеристиках на выборке из общего класса объектов (записей о качестве изделия);
- интерпретация полученных результатов на основе разработанных критериев;
- уточнение метамоделей, т.е. понятия объекта, его состояний, его характеристик, правил свёртки или критериев для интерпретации результатов.

Интеллектуализация процесса принятия решений хорошо укладывается в концепцию, предложенную В.В. Антоновым и К.А. Коневым [2021]. Общую схему сбора и анализа цифрового следа с использованием интеллектуальной подсистемы, построенной на основе методологии машинного обучения, можно представить схемой, показанной на рис. 4.

Таким образом, оцифровка источников информации о качестве изделия даёт возможность отслеживания его цифрового следа, что открывает возможности для поддержки принятия решений о качестве данной продукции с использованием интеллектуальных методов, таких как машинное обучение. Рассмотрим возможности для решения задачи определения

причины дефекта изделия авиаприборостроительного предприятия с использованием машинного обучения.

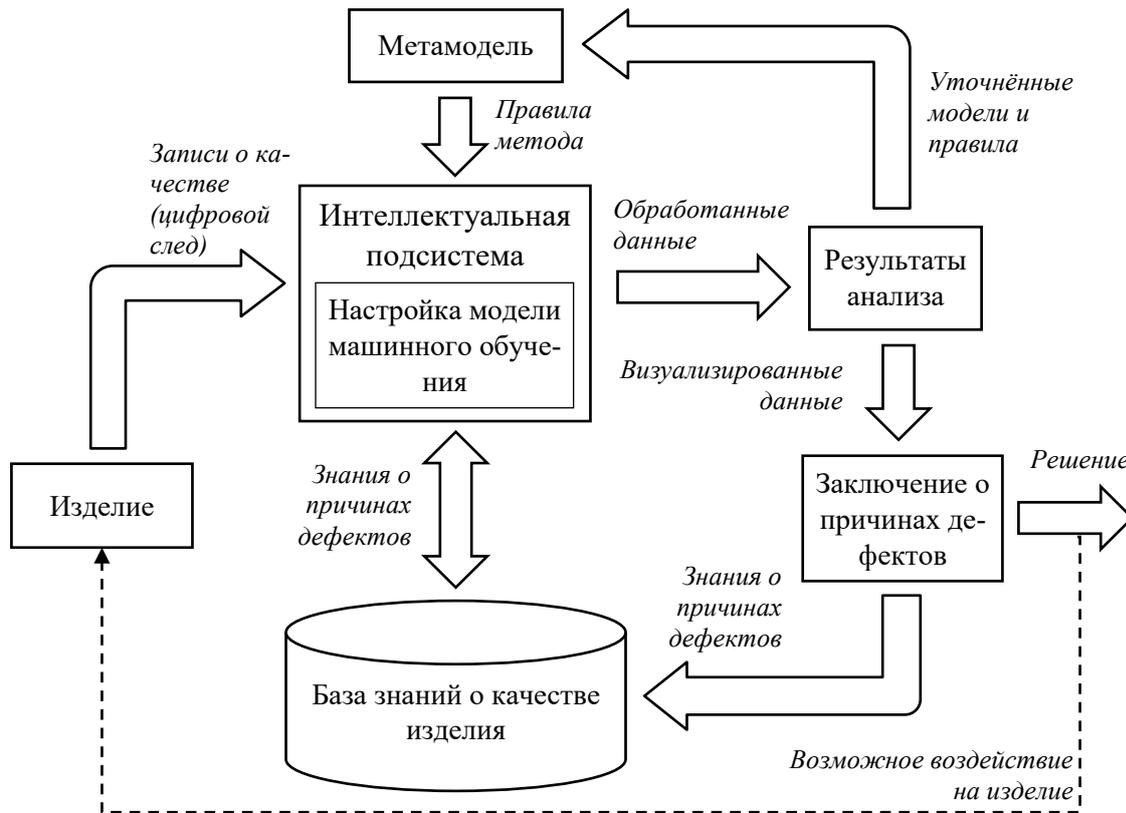


Рис. 4. Схема использования цифрового следа в интеллектуальной системе
 Fig. 4. Scheme of using a digital footprint in an intelligent system

Метод k -ближайших соседей машинного обучения для определения причины дефекта

Задача выбора причины дефекта является типичной задачей классификации в машинном обучении, связанной с анализом большого числа признаков, для решения которой воспользуемся методом k -ближайших соседей (k Nearest Neighbors, или kNN) [Толчеев, 2007, Madeh и др., 2020].

Как известно, метод k -ближайших соседей базируется на расчёте метрик и нормализации. Запишем алгоритм метода формально.

Допустим есть обучающая выборка $X = (x_i, y_i)_{i=1}^N$, где $x_i \in X$, $y_i \in Y$. Кроме того, задана симметричная по своим аргументам функция расстояния $\rho = X \otimes X \rightarrow [0, +\infty]$. Предположим, что требуется классифицировать новый объект u . С этой целью определим k наиболее близких к u в смысле расстояния ρ объектов обучающей выборки

$$X_k(u) = \{x_u^1, \dots, x_u^k\};$$

$$\forall x_{in} \in X_k(u) \forall x_{out} \in \frac{X}{X_k(u)}, \text{ причём } \rho(u, x_{in}) \leq \rho(u, x_{out})$$

Существуют как входящие в выборку элементы, так и не входящие в неё. Функция расстояния до входящих меньше или равна функции расстояния до не входящих элементов.

Метку класса объекта x_u^i обозначим y_u^i , тогда класс нового объекта определим как наиболее часто встречающийся класс среди объектов из $X_k(u)$:

$$a(u) = \operatorname{argmax}_{y \in Y} \sum_{i=1}^k I[y_u^i = y]$$

Выражение означает, что для каждой метки класса $y \in Y$ количество соседей u с такой меткой можно посчитать как сумму по всем соседям индикаторов событий, соответствующих тому, что метка соседа равна y .

Для расчёта ρ , как правило, используется обычное евклидово расстояние, которое рассчитывается следующим образом:

$$\rho(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2}$$

В некоторых случаях полезно использовать другие метрики: манхэттенскую метрику, Косинусное расстояние, метрику Минковского, Расстояние Жаккара и другие.

Разные атрибуты обычно обладают разным диапазоном представленных значений в выборке. Значения дистанции могут сильно зависеть от атрибутов с большими диапазонами, поэтому данные проходят через нормализацию.

MinMax-нормализация осуществляется следующим образом:

$$x' = \frac{x - \min[X]}{\max[X] - \min[X]}$$

в данном случае все значения будут находиться в диапазоне от 0 до 1.

Z-нормализация:

$$x' = \frac{x - M[X]}{\sigma[X]}$$

где σ – среднеквадратичное отклонение. В данном случае большинство значений попадает в диапазон.

Рассмотрим алгоритм реализации метода:

- Загрузка данных.
- Определение числа соседей k .
- Обход всего множества элементов и выполнение для каждого образца в данных:
 - определение расстояния между примером запроса и текущим примером из данных;
 - добавление индекса образца и его расстояние в упорядоченную коллекцию.
- Сортировка упорядоченной коллекции расстояний и индексов в порядке возрастания.
- Выбор первых k записей.
- Выбор метки выбранных k записей.
- Определение наиболее часто встречающегося значения выбранных ранее меток k .

Заметим, что не существует оптимального способа определить наилучшее значение для k , поэтому обычно оно определяется методом подбора. Но чаще всего наиболее предпочтительным значением для k является 5.

Применительно к задаче классификации причины дефекта осуществляется такой же выбор параметров из выборки по признаку наибольшей близости к оцениваемым свойствам. Например, если есть некоторый дефект, связанный с человеческим фактором, то согласно рисунку 2 и рассмотренной последовательности действий можно определить от-

носятся ли он к проблемам с квалификацией, мотивацией, способностями, производственным опытом или отношениями в коллективе. Для практической реализации достаточно сформировать оценочные значения классификационных признаков и пропустить их через алгоритм оценки методом k -ближайших соседей.

Рассмотрим пример реализации метода k -ближайших соседей с помощью специальных программных средств.

Построение модели классификации дефектов в Knime Analytics Platform

Knime Analytics Platform – инструмент для анализа данных с открытым кодом, предоставляющий возможности для решения широкого круга задач с использованием значительного числа методов по анализу данных, включая чтение данных из различных источников, преобразование и фильтрацию, собственно анализ, визуализацию и экспорт. В нём деятельность по заданию логики модели осуществляется через создание Workflow-модели, которая состоит из узлов, реализующих нужные функции, соединяющихся между собой стрелочками, показывающими направление движения данных [Knime Analytics Platform, 2017].

Для построения модели машинного обучения для решения задачи классификации причин дефектов сформируем признаки, по которым будет работать алгоритм. Для простоты выберем одну из ветвей на диаграмме Исикавы, связанную с оборудованием. Условимся считать, что для рассматриваемого бизнес-процесса пяти возможных классов проблем с оборудованием достаточно. Согласно рис.1 это проблемы: связанные с типом оборудования, связанные с износом отдельных элементов оборудования, связанные с обеспечением заданной точности оборудования, связанные с возрастом оборудования, связанные с качеством контроля условий эксплуатации.

Рассмотрим измеримые характеристики, которые будут влиять на возможность отнесения к классу одной из причин:

- наличие поломок в прошлом (качественная характеристика);
- наличие проблем с эксплуатацией оборудования в прошлом (качественная характеристика);
- возраст оборудования (в годах);
- отношение оставшегося времени до замены компонентов оборудования, подверженных ускоренному износу к интервалу между заменами (в долях);
- отношение оставшегося времени до замены расходных материалов в составе оборудования к интервалу между заменами (в долях);
- наличие проблем с точностью оборудования в прошлом (качественная характеристика);
- отношение оставшегося времени до проверки оборудования на точность к интервалу между проверками (в долях);
- соответствие квалификации обслуживающих оборудование лиц требованиям (качественная характеристика);
- наблюдаемый вид отказа оборудования (множество вариантов).

Если отобразить, например, 50 реальных отказов по причинам, связанным с оборудованием и для каждого из них определить перечисленные характеристики, то можно будет сформировать модель для обучения алгоритма на этой выборке по отбору тех причин, которые были выявлены экспертами. Далее на новых выборках можно будет использовать обученный алгоритм, который будет способен генерировать нужные решения (см. рис. 5).

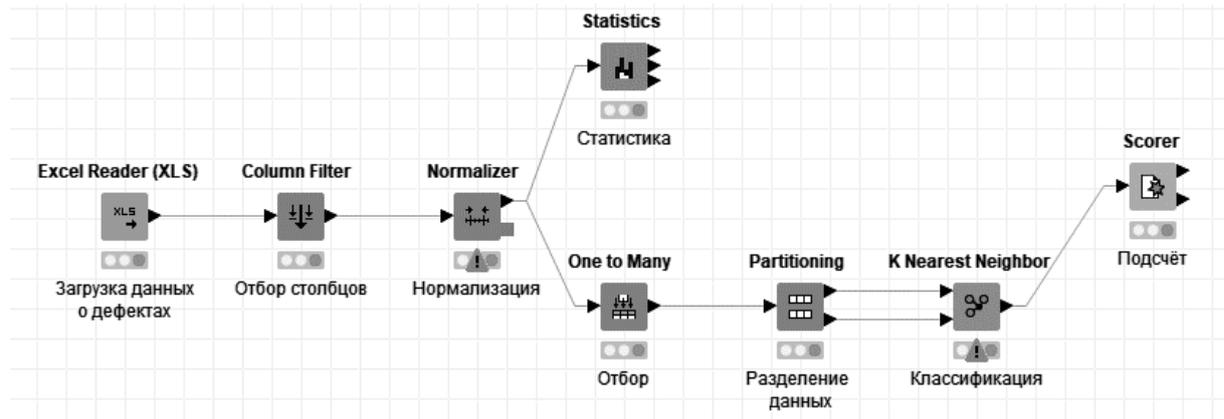


Рис. 5. Модель классификации машинного обучения для задачи выбора причины дефекта
Fig. 5. Machine Learning classification model for defect cause selection problem

При возникновении новых данных (причин дефектов) можно будет воспользоваться моделью для формирования новых кластеров (см. рис. 6).

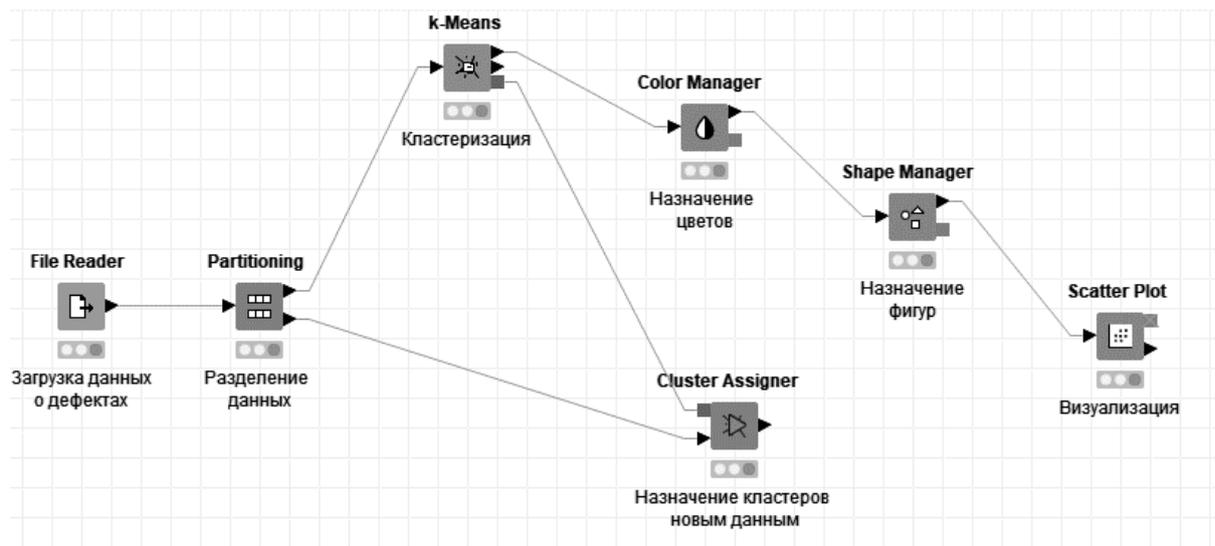


Рис. 6. Модель кластеризации машинного обучения для задачи выбора причины дефекта
Fig. 6. Machine Learning clustering model for defect cause selection problem

Таким образом, при помощи инструмента для анализа данных Knime Analytics Platform можно сформировать алгоритм машинного обучения в области обеспечения качества, что было продемонстрировано на примере задачи выбора причин дефектов оборудования на приборостроительном предприятии.

Заклучение

В результате проведённых исследований была проанализирована предметная область обеспечения качества, сформирована концептуальная модель интеллектуальной подсистемы системы поддержки принятия решений о качестве на основе алгоритма машинного обучения. Для задачи классификации причин дефектов проведено исследование предметной области, сформирован подход по формированию категорий причин на основе известного метода при-

чинно-следственной диаграммы Исикавы, который использован как метод формирования категорий для последующей цифровой обработки, построена структурная модель системы поддержки принятия решений о причинах дефектов. Проведён анализ проблемы сбора и подготовки информации о качестве на основе цифрового следа, т.е. так называемых записей о качестве, построена модель сбора для интеллектуальной подсистемы информации о качестве изделий, в данном случае представляющей собой описание характеристик среды изготовления изделия, которые могут быть выбраны в качестве категорий причин дефектов. В качестве метода машинного обучения предложен метод k-ближайших соседей, для применения которого был использован программный продукт Knime Analytics Platform.

Можно констатировать, что актуальная задача построения новой модели машинного обучения на основе сбора цифрового следа для задачи классификации причин дефектов предметной области обеспечения качества практически реализуема и может быть сведена к инженерной задаче применения, например, обладающей признаками научной новизны, методике сбора цифрового следа с категоризацией на основе причинно-следственной диаграммы Исикавы и использования инструментария Knime Analytics Platform.

Исследование проводилось для задач принятия решений в сфере обеспечения качества на предприятии авиационного приборостроения, но его результаты могут быть полезны и в других областях промышленности, в которых достаточно широкая номенклатура производимых изделий, деталей или сборочных единиц. В дальнейшем планируется апробация предложенных методов в других предметных областях.

Список источников

- Snow D. 2022. Machine Learning and Data Science Applications in Industry. URL: <https://github.com/firmai/industry-machine-learning> (accessed 19.09.2022).
Knime Analytics Platform – open source системы для анализа данных. 2017. URL: <https://habr.com/ru/post/320500/>. – (Дата обращения: 07.09.2022).

Список литературы

- Chen Yang, Weiming Shen, Xianbin Wang. 2018. The Internet of Things in Manufacturing: Key Issues and Potential Applications. *IEEE Systems, Man, and Cybernetics Magazine*. V. 4 (1): 6–15. DOI:10.1109/msmc.2017.2702391.
Kim C, Gupta R, Shah A, Madill E, Prabhu AV, Agarwal N. 2018. Digital Footprint of Neurological Surgeons. *World Neurosurg*, v. 113: e172–e178. doi: 10.1016/j.wneu.2018.01.210.
Madeh S. Piryonesi, Tamer E. El-Diraby. 2020. Role of Data Analytics in Infrastructure Asset Management: Overcoming Data Size and Quality Problems. *Journal of Transportation Engineering, Part B: Pavements*. 2020-06. V. 146, iss. 2. P. 04020022. ISSN 2573-5438 2573-5438, 2573-5438.
Negri Elisa. 2017. A review of the roles of Digital Twin in CPS-based production systems. *Procedia Manufacturing*. V. 11: 939–948. doi:10.1016/j.promfg.2017.07.198.
Zachman J. A. 1987. Framework for Information Systems Architecture. *IBM Systems Journal* 26: 276-292.
Антонов В.В., Конев К.А. 2021. Интеллектуальный метод поддержки принятия решений в типовой ситуации. *Онтология проектирования*. Т.11. 1(39): 126-136.
Антонов В.В., Конев К.А. Куликов Г.Г., Суворова В.А. 2021. Ситуационно-онтологическая методология принятия решений на примере бизнес-процессов авиаприборостроительного предприятия. *Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника*. Т. 21. 1: 102-115.
Исикава К. Японские методы управления качеством. 1988. М: «Экономика», 199 с.
Коробеев А.И., Чучаев А.И. 2018. Беспилотные транспортные средства, оснащённые системами искусственного интеллекта: проблемы правового регулирования. *Азиатско-тихоокеанский регион: экономика, политика, право*. Т. 20. 3: 117-132. – DOI 10.24866/1813-3274/2018-3/117-132
Толчеев В.О. 2007. Модифицированный и обобщенный метод ближайшего соседа для классификации библиографических текстовых документов. *Заводская лаборатория. Диагностика материалов*. Т. 75. 7: 63-70.
Флах П. Машинное обучение. 2015. М., ДМК Пресс, 400 с.

References

- Chen Yang, Weiming Shen, Xianbin Wang. 2018. The Internet of Things in Manufacturing: Key Issues and Potential Applications // *IEEE Systems, Man, and Cybernetics Magazine*. V. 4 (1). pp. 6–15. ISSN 2333-942X. doi:10.1109/msmc.2017.2702391.
- Kim C, Gupta R, Shah A, Madill E, Prabhu AV, Agarwal N. 2018. Digital Footprint of Neurological Surgeons. *World Neurosurg*, v. 113, pp. e172–e178. doi: 10.1016/j.wneu.2018.01.210.
- Madeh S. Piryonesi, Tamer E. El-Diraby. 2020. Role of Data Analytics in Infrastructure Asset Management: Overcoming Data Size and Quality Problems. *Journal of Transportation Engineering, Part B: Pavements*. 2020-06. V. 146, iss. 2. P. 04020022. ISSN 2573-5438 2573-5438, 2573-5438.
- Negri Elisa. A review of the roles of Digital Twin in CPS-based production systems. *Procedia Manufacturing*. V. 11, 2017: 939–948. doi:10.1016/j.promfg.2017.07.198.
- Overview of Knime Analytics Platform – an open source system for data analysis. 2017. [Electronic resource] URL: <https://habr.com/ru/post/320500/> (accessed 17.09.2022)
- Snow D. 2022. Machine Learning and Data Science Applications in Industry. [Electronic resource] URL: <https://github.com/firmai/industry-machine-learning> (accessed 19.09.2022).
- Zachman J. A. 1987. Framework for Information Systems Architecture. *IBM Systems Journal* 26: 276-292.
- Antonov V.V., Konev K.A. 2021. Intelligent decision support method in a typical situation. *Design ontology*. V.11. 1(39): 126-136. (in Russian).
- Antonov V.V., Konev K.A. Kulikov G.G., Suvorova V.A. 2021. Situational-ontological decision-making methodology on the example of business processes of an aircraft instrument-making enterprise. *Bulletin of the South Ural State University. Series: Computer technologies, control, radio electronics*. V. 21. 1: 102-115. (in Russian).
- Ishikawa K. 1988. Japanese methods of quality management, Moscow. Economics, 199 p. (in Russian)
- Korobeev A.I., Chuchaev A.I. 2018. Unmanned vehicles equipped with artificial intelligence systems: problems of legal regulation // *Asia-Pacific region: economics, politics, law*. V. 20. 3: 117-132. – DOI 10.24866/1813-3274/2018-3/117-132
- Tolcheev V.O. 2007. Modified and generalized nearest neighbor method for classification of bibliographic text documents, in *Zavodskaya lab. Material diagnostics*. V. 75. no. 7. pp. 63-70, (in Russian).
- Flach P. Machine learning, 2015. Moscow: DMK Press, 400 p. (in Russian).

Конфликт интересов: о потенциальном конфликте интересов не сообщалось.

Conflict of interest: no potential conflict of interest related to this article was reported.

ИНФОРМАЦИЯ ОБ АВТОРЕ

Конеv Константин Анатольевич, к.т.н., доцент, доцент кафедры автоматизированных систем управления Уфимского университета науки и технологий (быв. УГАТУ), г. Уфа, Россия

INFORMATION ABOUT THE AUTHOR

Konstantin A. Konev, Candidate of Technical Sciences, Associate Professor, Associate Professor Department of Automated Control Systems, Ufa University of Science and Technology, Ufa, Russia