

(19)



Евразийское  
патентное  
ведомство

(21) 202392893

(13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки  
2024.04.25

(51) Int. Cl. G06N 3/08 (2023.01)  
G06N 7/04 (2006.01)  
G06N 20/00 (2019.01)  
G05B 13/04 (2006.01)

(22) Дата подачи заявки  
2023.11.13

(54) УСТРОЙСТВО ДЛЯ СОЗДАНИЯ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ НА ОСНОВЕ ДАННЫХ

(31) 2022130782

(32) 2022.11.25

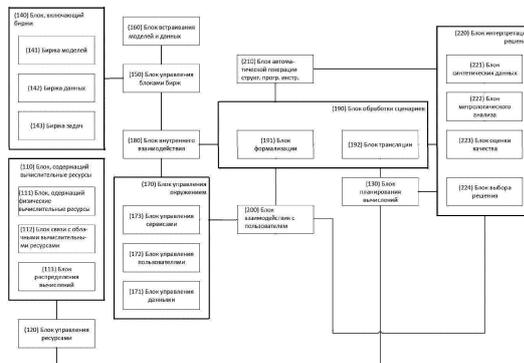
(33) RU

(71) Заявитель:  
ФЕДЕРАЛЬНОЕ  
ГОСУДАРСТВЕННОЕ  
АВТОНОМНОЕ  
ОБРАЗОВАТЕЛЬНОЕ  
УЧРЕЖДЕНИЕ ВЫСШЕГО  
ОБРАЗОВАНИЯ  
"НАЦИОНАЛЬНЫЙ  
ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ  
ИТМО" (УНИВЕРСИТЕТ ИТМО)  
(RU)

(72) Изобретатель:  
Бухановский Александр Валерьевич,  
Калюжная Анна Владимировна,  
Насонов Денис Александрович,  
Иванов Сергей Владимирович (RU)

(74) Представитель:  
Абраменко О.И. (RU)

(57) Изобретение предназначено для автоматической генерации структуры и исполнимого кода модели машинного обучения, обеспечивающего решение задачи пользователя с заданной точностью и обучаемой на распределённых вычислительных ресурсах за заданное время. Технический результат изобретения заключается в формировании модели машинного обучения (в форме исполнимого кода). Изобретение было создано в рамках договора о предоставлении средств юридическому лицу, индивидуальному предпринимателю на безвозмездной и безвозвратной основе в форме гранта, источником финансового обеспечения которых полностью или частично является субсидия, предоставленная из федерального бюджета, № 70-2021-00141 от 2 ноября 2021 по мероприятию 5.2 Техническое проектирование технологического полигона для испытаний и оценки качества систем с элементами сильного ИИ в промышленности, направление развития ИИ - генеративный искусственный интеллект.



A1

202392893

202392893

A1

## **Устройство для создания модели машинного обучения на основе данных**

### **Область техники**

Техническое решение относится к области вычислительной техники, а более конкретно к области организации вычислительных устройств и устройств хранения данных для решения задач проектирования, обучения и эксплуатации цифровых моделей, построенных методами машинного обучения.

### **Уровень техники**

Бурное развитие компьютерных технологий в последнее десятилетие, широкое распространение разнообразных вычислительных устройств (персональных компьютеров, ноутбуков, смартфонов и т.д.), а также обеспечение круглосуточного доступа к данным из любой точки мира стали мощным стимулом для осуществления глобальной компьютеризации в разнообразных сферах деятельности и огромном количестве пользовательских задач (от интернет-сёрфинга до работы с облачными данными на удалённых системах и электронного документооборота).

Существенное развитие получила и область машинного обучения – класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи на основе заранее известной модели, а создание моделей, которые обучаются за счёт применения решений множества сходных задач. Для построения таких моделей используются средства математической статистики, численных методов, методов оптимизации, теории вероятностей, теории графов, различные техники работы с данными в цифровой форме.

В настоящий момент существует ряд средств (платформ) разработки цифровых моделей машинного обучения на основе стандартных компонентов и библиотек, обеспечивающими процесс их обучения на облачных вычислительных ресурсах. Основные отличия между ними заключаются в трёх качествах:

- способе представления моделей (непосредственное кодирование на языке низкого уровня, работа с шаблонами, low-code- и no-code диаграммами);
- организации доступа к вычислительным ресурсам (виртуальные машины, физические ресурсы, иные распределённые среды) и управления ими;
- способе разработки моделей (ручной или автоматизированный, с использованием механизмов автоматического машинного обучения).

Одни платформы, такие как ML Space, DataSphere и Google Colab, реализуют базовые функции для ручной разработки моделей машинного обучения в виде кодов на языке Python и осуществляют управление вычислительными ресурсами в облаке без учёта структуры самой модели (т.е. не обеспечивается соответствие выбора вычислительных ресурсов составу решаемых задач). Другие платформы, в том числе Microsoft Azure, реализуют упрощённые функции ручной разработки (в виде low-code-диаграмм), но не обеспечивают прозрачности структуры модели и использования ресурсов разной архитектуры. Третьи платформы, в том числе Knime, также реализуют развитые механизмы ручной разработки с помощью low-code-диаграмм, однако не управляют вычислительными ресурсами для исполнения моделей (исполняется на локальном компьютере).

Настоящее техническое решение позволяет решать как задачи задания общей структуры модели как при помощи упрощенных функций ручной разработки в виде визуальной последовательности команд (low-code-диаграммы), так и на языке низкого уровня, с последующей автоматической трансляцией в детальную исполнимую форму с помощью методов автоматического машинного обучения, как и задачи эффективного планирования использования вычислительных ресурсов, доступных платформе, непосредственно на этапе трансляции, за счёт управления структурой модели машинного обучения.

### **Сущность**

Устройство создания модели машинного обучения на основе данных, которое содержит соединенные между собой при помощи электрических неразъемных соединений (с обеспечением возможности приема/передачи данных) блоки:

а) блок взаимодействия с пользователем, выполненный с возможностью обеспечения интерфейса человеко-компьютерного взаимодействия, содержащего средства для конструирования, разработки и исполнения композитных приложений, мониторинга вычислительных процессов и формирования отчётов на основе выполненных расчётов;

приема и передачи информации в блок управления окружением, блок обработки сценариев 190 и блок интерпретации решений;

б) блок управления окружением, выполненный с возможностью регистрации и управления правами пользователей;

управления пользовательскими данными и сервисами, обеспечивающими пользователям доступ для работы с устройством;

приема и передачи информации в блок внутреннего взаимодействия и блок взаимодействия с пользователем;

в) блок бирж данных, содержащий информацию о биржах, предназначенных для каталогизации моделей, данных и задач, реализуемых пользователями системы и представленных в унифицированной форме, обеспечивающей их использование пользователями в соответствии с их правами, и выполненный с возможностью приема и передачи информации в блок управления блоками бирж;

г) блок управления блоками бирж, выполненный с возможностью обеспечения каталогизации и структурирования задач, данных и моделей, созданных пользователями и хранящихся в биржах блока хранения данных бирж, приема и передачи информации в блок встраивания моделей и данных и блок внутренних взаимодействий;

д) блок обработки сценариев, выполненный с возможностью анализа формального описания задачи пользователя на создание модели машинного обучения, поступающей через блок взаимодействия с пользователем, автоматического преобразование полученной задачи в исполнимую форму программных инструкций для обучения модели посредством блока автоматической генерации структуры программных инструкций, выполненный с возможностью приема и передачи информации в блок внутренних взаимодействий, блок взаимодействия с пользователем, блок планирования вычислений и блок интерпретации решений;

е) блок автоматической генерации структуры программных инструкций, соответствующего заданным пользователем критериям, выполненный с возможностью приема и передачи информации в блок обработки сценариев и блок интерпретации решения;

ж) блок планирования вычислений, выполненный с возможностью построения оптимального расписания исполнения всех элементов программных инструкций таким образом, чтобы обеспечить выполнение

критериев пользователя при минимальной загрузке вычислительных ресурсов, приема и передачи информации в блок управления ресурсами и блок обработки сценариев;

з) блок вычислений, содержащий вычислительные ресурсы, обеспечивающие распределённое выполнение расчётов в целях обучения модели, посредством запуска программных инструкций, сформированного блоком обработки сценариев в соответствии с расписанием блока планирования вычислений, выполненный с возможностью приема и передачи информации в блок управления ресурсами;

и) блок управления ресурсами, выполненный с возможностью обеспечения загрузки и исполнения расчётных задач для обучения и использования моделей, в форме программных инструкций на вычислительных ресурсах, приема и передачи информации в блок вычислений, содержащий вычислительные ресурсы и блок планирования вычислений;

к) блок встраивания моделей и данных, выполненный с возможностью размещения в блоке управления блоками бирж новых данных и моделей, созданных вне данного устройства, выполненный с возможностью приема и передачи информации в блок управления блоками бирж;

л) блок интерпретации решений, выполненный с возможностью обеспечения контроля качества обученных моделей путём оценки их точности по сравнению с эталонными решениями, полученными с помощью блока автоматической генерации структуры программных инструкций на синтетических массивах данных, приема и передачи информации в блок автоматической генерации структуры программных инструкций, блок обработки сценариев, блок планирования вычислений и блок взаимодействия с пользователем;

м) блок внутренних взаимодействий, выполненный с возможностью организации совместной работы всех внутренних блоков и маршрутизации потоков данных в устройстве, приема и передачи информации в блок управления блоками бирж, блок управления окружением и блок обработки сценариев;

### **Раскрытие технического решения**

Техническое решение предназначено для автоматической генерации структуры и исполнимого кода модели машинного обучения, обеспечивающего решение задачи пользователя с заданной точностью и обучаемой на распределённых вычислительных ресурсах за заданное время.

Технический результат настоящего технического решения заключается в формировании модели машинного обучения (в форме исполнимого кода).

Данный технический результат достигается с помощью использования устройства создания модели машинного обучения на основе данных, которая содержит: блок взаимодействия с пользователем, предназначенный для обеспечения интерфейса человеко-компьютерного взаимодействия, включая средства для конструирования, разработки и исполнения программных инструкций, мониторинга вычислительных процессов и формирования отчётов на основе выполненных расчётов; блок управления окружением, предназначенный для обеспечения внутрисистемных функций устройства в части регистрации и управления правами пользователей, управления пользовательскими данными, а также сервисами, обеспечивающими пользователям доступ к функциональности устройства; блок, включающий биржи, предназначенные для каталогизации моделей, данных и задач, реализуемых разными пользователями устройства и представленных в унифицированной форме, обеспечивающей их повторное использование другими пользователями (кроме владельца) в соответствии с их правами; блок управления блоками бирж, предназначенный для обеспечения каталогизации

и структурирования задач, данных и моделей, созданных разными пользователями и хранящихся в соответствующих биржах; блок обработки сценариев, предназначенный для обеспечения анализа формального описания задачи пользователя на создание модели машинного обучения с последующим его автоматическим преобразованием в исполнимую форму программной инструкции для обучения модели посредством блока автоматической генерации структуры программных инструкций; блок автоматической генерации структуры программных инструкций, предназначенный для обеспечения автоматической генерации структуры программных инструкций, соответствующих заданным пользователем критериям; блок планирования вычислений, предназначенный для построения оптимального расписания исполнения всех элементов программных инструкций таким образом, чтобы обеспечить выполнение критериев пользователя при минимальной загрузке вычислительных ресурсов; вычислительные ресурсы, предназначенные для распределённого выполнения расчётов в целях обучения модели, посредством запуска программных инструкций, сформированного блоком обработки сценариев в соответствии с расписанием блока планирования вычислений; блок управления ресурсами, предназначенный для обеспечения загрузки и исполнения расчётных задач для обучения и использования моделей, в форме программных инструкций на вычислительных ресурсах; блок встраивания моделей и данных, предназначенный для размещения в биржах новых данных и моделей, созданных вне данного устройства и загруженных пользователем; блок интерпретации решений, предназначенный для обеспечения контроля качества обученных моделей путём оценки их точности по сравнению с эталонными решениями, полученными с помощью блока автоматической генерации структуры программных инструкций на синтетических массивах данных; блок внутреннего взаимодействия, предназначенный для организации

совместной работы всех внутренних блоков и маршрутизации потоков данных в системе.

В другом частном случае реализации устройства блок, содержащий вычислительные ресурсы содержит по меньшей мере: блок, содержащий физические вычислительные ресурсы, представляющие собой высокопроизводительные комплексы с собственной системой управления, предназначенные для реализации ресурсоёмких расчётов для обучения моделей, распараллеливаемые с учётом специфики их архитектуры; блок связи с облачными вычислительными ресурсами, который обеспечивает возможность взаимодействия с по меньшей мере публичными, частными облаками, предназначенными для обеспечения использования обученных моделей как сервисов, а также для хранения промежуточных данных расчётов; блок распределённых вычислений, предназначенный для обеспечения доступа к распределённо хранящимся данным под единой системой управления на основе по меньшей мере пакетной, потоковой модели.

Ещё в одном частном случае реализации устройства программные инструкции организуются посредством контейнеров.

В другом частном случае реализации устройства блок управления окружением содержит по меньшей мере: блок управления данными, предназначенный для обеспечения индексации данных в соответствии с заданной структурной моделью, а также для распределения по различным вычислительным ресурсам; блок управления пользователями, предназначенный для регистрирования в системе новых пользователей, установке прав пользователей, доступа к данным, задачам и ресурсам, а также для определения статистики их использования; блок управления сервисами, предназначенный для обеспечения доступа пользователя к данным, моделям и задачам как к цифровым сущностям, отображаемым в среде разработки.

Ещё в одном частном случае реализации устройства/ блок взаимодействия с пользователем содержит по меньшей мере: редактор исходного кода, предназначенный для формирования и проверки исходного кода задачи на языке Python; редактор визуальной последовательности команд (low-code-диаграмм), предназначенный для формирования описания исходной задачи в виде графа, и дальнейшей трансляции его к текстовому описанию на языке Python; среду разработки, предназначенную для управления процессами загрузки и подготовки данных на основании команд от пользователя, создания и запуска на исполнение расчётных задач, а также визуализации и анализа результатов обучения и применения моделей; блок генерации отчётов, предназначенный для визуализации результатов расчётов в виде таблиц, графиков и иных видов изображений, а также формирования интегрального описания решённой задачи по созданию и обучению модели на данных.

В другом частном случае реализации устройства блок, включающий биржи содержит по меньшей мере: биржу моделей, которая содержит модели машинного обучения, созданные разными пользователями, в унифицированном представлении, и делает доступными их другим пользователям (кроме владельца) в соответствии с их правами; биржу данных, которая содержит входные, выходные и синтетические данные, сформированные разными пользователями, в унифицированном представлении, и делает доступными их другим пользователям (кроме владельца) в соответствии с их правами; биржу задач, которая содержит описания прикладных задач машинного обучения, содержащих одну или несколько моделей, а также соответствующие им данные, при этом задачи формируются разными пользователями, хранятся в унифицированном представлении, и доступны другим пользователям (кроме владельца) в соответствии с их правами.

Ещё в одном частном случае реализации устройства блок интерпретации решений содержит по меньшей мере: блок синтетических данных, предназначенный для обеспечения посредством вероятностных моделей и бутстреппирования генерации массивов данных, обладающих теми же свойствами, что и исходные, для формирования условий проверки качества работы обученных моделей; блок метрологического анализа, реализованный на основе синтетических данных с помощью функций блока автоматической генерации структуры программных инструкций и предназначенный для постройки эталонных моделей, которые предназначены для сопоставления с моделями пользователя при оценке их точности и устойчивости; блок оценки качества, предназначенный для оценки характеристики точности (по набору гостированных метрик) и устойчивости моделей, разработанных и обученных пользователем, с использованием синтетических данных; блок выбора решения, предназначенный для ранжирования различных реализации моделей, созданных пользователем, с точки зрения соответствия многомерным критериям качества.

В другом частном случае реализации устройства блок обработки сценариев содержит по меньшей мере: блок формализации, предназначенный для обеспечения высокоуровневой формализации и выполнения задачи пользователя по подготовке входных и выходных данных, формированию структуры модели в виде программных инструкций и запуску модели на обучение; блок трансляции, предназначенный для обеспечения трансляции, где трансляция представляет собой перевод высокоуровневого представления программных инструкций задачи в виде графа в детализированную исполнимую форму с определением реализаций всех расчётных функций.

Ещё в одном частном случае реализации устройства в блоке, включающем биржи, размещают модели по меньшей мере в форме: исходных кодов; откомпилированных библиотек.

## Краткое описание чертежей

На Фиг. 1 изображено устройство создания модели машинного обучения на основе данных.

### Описание вариантов осуществления технического решения

Объекты и признаки настоящего технического решения, способы для достижения этих объектов и признаков станут очевидными посредством отсылки к примерным вариантам осуществления. Однако настоящее техническое решение не ограничивается примерными вариантами осуществления, раскрытыми ниже, оно может воплощаться в различных видах. Сущность, приведённая в описании, является ничем иным, как конкретными деталями, необходимыми для помощи специалисту в области техники в исчерпывающем понимании технического решения, и настоящее техническое решение определяется в объёме приложенной формулы.

Введём ряд определений и понятий, которые будут использоваться при описании вариантов осуществления изобретения.

*Визуальная последовательность команд (Low-code-диаграмма)* – способ представления программы в виде графа, вершинами которого являются отдельные вычислительные операции, а связи регламентируют операции с данными. При этом кодирование на языке низкого уровня связано только с отдельными операциями (тонкой настройкой программы).

*Автоматическое машинное обучение* (англ. AutoML) – область технологий искусственного интеллекта, позволяющая автоматически определять структуру и параметры моделей машинного обучения на

конкретных наборах данных в соответствии с критериями качества, задаваемыми пользователем.

*Бутстрепирование* – способ создания синтетических массивов данных, сохраняющих вероятностные свойства исходной выборки, путём комбинаторных операций над ее членами.

*Индексация данных* – априорное присваивание данным индексов в соответствии с предполагаемым способом их использования с целью ускорения операций поиска и обработки в ходе работы устройства.

*Программная инструкция* – часть композитного приложения, исполняющаяся непосредственно в блоке устройства, позволяющая составленному из нескольких логически взаимосвязанных компонентов (отдельных приложений) композитному приложению, исполняемому на распределённых вычислительных ресурсах, представимому в форме потока работ (workflow), осуществлять обмен данными с устройством.

*Контейнер* – структура данных, позволяющая инкапсулировать в себе объекты разных типов (включая формальные описания моделей и пользовательские данные), обеспечивающая удобное управление ими в части исполнения на разных вычислительных ресурсах, масштабирования и обновления.

*Критерии* — требования достижения номинальных характеристик, связанных с:

- точностью модели: метрики классификации (точность, логарифмическая потеря, площадь под кривой ROC и другие); метрики для прогнозирования (матрица путаницы, классификационный отчёт и другие), показатели регрессии (средняя

абсолютная ошибка, средняя квадратичная ошибка, коэффициент детерминации модели, и другие);

- временными и ресурсными характеристиками, необходимыми для обучения и использования модели (время вычислений, количество вычислительных ресурсов (процессоров, ядер, объем памяти), параллельное ускорение, параллельная эффективность, реактивность и прочее).

*Модель машинного обучения* – зависимость (соотношение) между данными, построенное путём обобщения априорного опыта и данных с помощью определённого алгоритма искусственного интеллекта.

*Планирование использования вычислительных ресурсов* – составление расписаний и условий для запуска каждого из компонентов программных инструкций на отдельных вычислительных ресурсах распределённой среды таким образом, чтобы обеспечить достижение заданного показателя эффективности (например, минимального времени расчётов).

*Структура модели машинного обучения* – перечень основных функциональных операций над данными, а также перечень условий передачи и получения данных между этими операциями. Может быть описана в форме визуальной последовательности команд (low-code-диаграммы) и реализована в форме программных инструкций.

**Фиг. 1** представляет устройство создания модели машинного обучения на основе данных.

Устройство создания модели машинного обучения на основе данных, содержащая по меньшей мере один постоянный машиночитаемый носитель, включающий взаимодействующие между собой средства, содержит: блок, содержащий вычислительные ресурсы **110**, блок управления ресурсами **120**,

блок планирования вычислений **130**, блок управления окружением **170**, блок внутреннего взаимодействия **180**, блок обработки сценариев **190**, блок, включающий биржи **140**, блок управления блоками бирж **150**, блок встраивания моделей и данных **160**, блок интерпретации решений **220**, блок автоматической генерации структуры программных инструкций **210**.

Описываемое устройство создания модели машинного обучения на основе данных представляет собой физические блоки, соединенные между собой и обрабатывающие особым образом данные.

Блок, содержащий вычислительные ресурсы **110** предназначен для распределённого выполнения расчётов в целях обучения модели, посредством запуска программных инструкций, сформированного блоком обработки сценариев **190** в соответствии с расписанием блока планирования вычислений **130**.

Блок, содержащий вычислительные ресурсы **110** выполнен с возможностью приёма/передачи данных из блока управления ресурсами **120**.

В одном из вариантов реализации устройства блок, содержащий вычислительные ресурсы **110** содержит по меньшей мере:

- блок, содержащий физические вычислительные ресурсы **111**, представляющие собой высокопроизводительные комплексы с собственной системой управления, предназначенные для реализации ресурсоёмких расчётов для обучения моделей, распараллеливаемые с учётом специфики их архитектуры;
- блок связи с облачными вычислительными ресурсами **112**, который обеспечивает возможность взаимодействия по меньшей мере с публичными, частными облаками, которые

предназначены для обеспечения использования обученных моделей как сервисов

(SaaS), а также для хранения промежуточных данных расчётов (DaaS);

- блок распределённых вычислений **113**, предназначенный для обеспечения доступа к распределённо хранящимся данным под единой системой управления на основе по меньшей мере пакетной, потоковой модели.

Блок, содержащий физические вычислительные ресурсы **111**, блок связи с облачными вычислительными ресурсами **112**, а также блок распределённых вычислений **113** выполнены с возможностью приёма-передачи данных из блока управления ресурсами **120**.

Ещё в одном из вариантов реализации устройства блок, содержащий физические вычислительные ресурсы **111** реализован по меньшей мере двумя классами вычислительных систем:

- гибридными вычислительными комплексами для машинного обучения NVIDIA DGX-100 и DGX-1 на основе графических ускорителей;
- вычислительным кластером для обработки данных (31 сервер CISCO UCS C 220 M5, коммуникационная сеть на основе CISCO Nexus 9300).

Ещё в одном из вариантов реализации устройства блок связи с облачными вычислительными ресурсами **112** является блоком, обеспечивающим доступ к публичным облакам, включая Amazon.

Блок управления ресурсами **120** предназначен для обеспечения загрузки и исполнения расчётных задач для обучения и использования моделей, в форме программных инструкций на вычислительных ресурсах **110**.

Блок управления ресурсами **120** выполнен с возможностью приёма/передачи данных из вычислительных ресурсов **110** и блока планирования вычислений **130**.

Блок планирования вычислений **130** предназначен для построения оптимального расписания исполнения всех элементов программных инструкций таким образом, чтобы обеспечить выполнение критериев пользователя при минимальной загрузке вычислительных ресурсов **110**.

Блок планирования вычислений **130** выполнен с возможностью приёма/передачи данных из блока управления ресурсами **120** и блока обработки сценариев **190** (в частности блока трансляции **192**).

В одном из вариантов реализации устройства программные инструкции организовывают посредством контейнеров (например, средствами виртуализации).

Блок, включающий биржи **140**, в котором биржи предназначены для каталогизации моделей, данных и задач, реализуемых разными пользователями устройства и представленных в унифицированной форме, обеспечивающей их повторное использование другими пользователями (кроме владельца) в соответствии с их правами.

Блок, включающий биржи **140** выполнены с возможностью приёма-передачи данных из блока управления блоками бирж **150**.

В одном из вариантов реализации устройства в блоке, включающем биржи данных **140**, биржи данных являются логическими конструкциями,

однако физически они работают на ресурсах блока распределённых вычислений **113**.

Ещё в одном из вариантов реализации устройства блок, включающий биржи **140** содержит по меньшей мере:

- биржу моделей (блок биржи моделей) **141**, которая содержит модели машинного обучения, созданные разными пользователями, в унифицированном представлении, и делает доступными их другим пользователям (кроме владельца) в соответствии с их правами;
- биржу данных (блок биржи данных) **142**, которая содержит входные, выходные и синтетические данные, сформированные разными пользователями, в унифицированном представлении, и делает доступными их другим пользователям (кроме владельца) в соответствии с их правами;
- биржу задач (блок биржи задач) **143**, которая содержит описания прикладных задач машинного обучения, содержащих одну или несколько моделей, а также соответствующие им данные, при этом задачи формируются разными пользователями, хранятся в унифицированном представлении, и доступны другим пользователям (кроме владельца) в соответствии с их правами.

Ещё в одном из вариантов реализации устройства в блоке, включающем биржи **140** размещают модели по меньшей мере в форме:

- исходных кодов;
- откомпилированных библиотек.

Блок управления блоками бирж **150** предназначен для обеспечения каталогизации и структурирования задач, данных и моделей, созданных разными пользователями и хранящихся в соответствующих биржах **140**.

Блок управления блоками бирж **150** выполнен с возможностью приёма-передачи данных из блока, включающего биржи **140**, блока встраивания моделей и данных **160** и блока внутренних взаимодействий **180**.

Блок встраивания моделей и данных **160** предназначен для размещения в блоке, включающем биржи **140** новых данных и моделей, созданных вне данного устройства и загруженных пользователем.

Блок встраивания моделей и данных **160** выполнен с возможностью приёма-передачи данных из блока управления блоками бирж **150**.

Блок управления окружением **170** обеспечивает внутренние функции устройства в части регистрации и управления правами пользователей, управления пользовательскими данными, а также сервисами, обеспечивающими пользователям доступ к функциональностям устройства.

Блок управления окружением **170** выполнен с возможностью приёма-передачи данных из блока внутреннего взаимодействия **180** и блока взаимодействия с пользователем **200**.

В одном из вариантов реализации устройства блок управления окружением **170** содержит по меньшей мере:

- блок управления данными **171**, предназначенный для обеспечения индексации данных в соответствии с заданной структурной моделью, а также для распределения по вычислительным ресурсам **110**;

- блок управления пользователями **172**, предназначенный для регистрации в системе новых пользователей, установке прав пользователей, доступа к данным, задачам и вычислительным ресурсам **110**, а также для определения статистики их использования;
- блок управления сервисами **173**, предназначенный для обеспечения доступа пользователя к данным, моделям и задачам как к цифровым сущностям, отображаемым в блоке взаимодействия с пользователем **200**.

Блок управления сервисами **173** выполнен с возможностью приёма/передачи данных из блока взаимодействия с пользователем **200**.

Блок внутреннего взаимодействия **180** предназначен для организации совместной работы всех внутренних блоков и маршрутизации потоков данных в системе.

Блок внутреннего взаимодействия **180** выполнен с возможностью приёма-передачи данных из блока управления блоками бирж **150**, блока управления окружением **170** и блока обработки сценариев **190** (в частности блока формализации **191**).

Блок обработки сценариев **190** является ключевым элементом устройства и обеспечивает анализ формального описания задачи пользователя с последующим его автоматическим преобразованием в исполнимую форму программных инструкций для обучения модели посредством блока автоматической генерации структуры программных инструкций **210**.

Блок обработки сценариев **190** выполнен с возможностью приёма-передачи данных из блока внутренних взаимодействий **180**, блок

взаимодействия с пользователем **200** (в частности через блок формализации **191**), блоком планирования вычислений **130** (в частности через блок трансляции **192**) и блоком интерпретации решения **220** (в частности через блок трансляции **192**).

В одном из вариантов реализации устройства блок обработки сценариев **190** содержит по меньшей мере:

- блок формализации **191**, предназначенный для обеспечения высокоуровневой формализации и выполнения задачи пользователя по подготовке входных и выходных данных, формированию структуры модели в виде программных инструкций и запуску модели на обучение;
- блок трансляции **192**, предназначенный для обеспечения трансляции, где трансляция представляет собой перевод высокоуровневого представления программных инструкций задачи в виде графа в детализированную исполнимую форму с определением реализаций всех расчётных функций.

Блок формализации **191** выполнен с возможностью приёма-передачи данных из блока внутреннего взаимодействия **180**, блока трансляции **192**, блока автоматической генерации структуры программных инструкций **210** и блока взаимодействия с пользователем **200**.

Блок трансляции **192** выполнен с возможностью приёма-передачи данных из блока формализации **191**, блока планирования вычислений **130** и блока интерпретации решения **220** (в частности, блока оценки качества **223**).

Блок взаимодействия с пользователем **200** предназначен для обеспечения интерфейса человеко-компьютерного взаимодействия, реализуемого средой разработки, включенной в блок взаимодействия с

пользователем, которая включает средства для конструирования, разработки и исполнения программных инструкций, мониторинга вычислительных процессов и формирования отчётов на основе выполненных расчётов.

Блок взаимодействия с пользователем **200** выполнен с возможностью приёма-передачи данных из блока управления окружением **170** (в частности блока управления сервисами **173**), блока обработки сценариев **190** (в частности блока формализации **191**) и блока интерпретации решения **220** (в частности блока выбора решения **224**).

Блок автоматической генерации структуры программных инструкций (автоМЛ) **210** предназначен для обеспечения автоматической генерации структуры программных инструкций, соответствующего заданным пользователем критериям.

Блок автоматической генерации структуры программных инструкций **210** выполнен с возможностью приёма-передачи данных из блока обработки сценариев **190** (в частности, блока формализации **191**) и блока интерпретации решения **220** (в частности блока синтетических данных **221**).

Блок интерпретации решений **220** обеспечивает контроль качества обученных моделей путём оценки их точности по сравнению с эталонными решениями, полученными с помощью блока автоматической генерации структуры программных инструкций **210** на синтетических массивах данных.

Блок интерпретации решений **220** выполнен с возможностью приёма/передачи данных из блока автоматической генерации структуры программных инструкций **210**, блока обработки сценариев **190** (в частности блока трансляции **192**), блока планирования вычислений **130** и блока взаимодействия с пользователем **200**.

В одном из вариантов реализации устройства блок интерпретации решений **220** содержит по меньшей мере:

- блок синтетических данных **221**, предназначенный для обеспечения посредством вероятностных моделей и бутстрепирования генерации массивов данных, обладающих теми же свойствами, что и исходные, для формирования условий проверки качества работы обученных моделей;
- блок метрологического анализа **222**, реализованный на основе синтетических данных с помощью функций блока автоматической генерации структуры программных инструкций **210** и предназначенный для постройки эталонных моделей, которые предназначены для сопоставления с моделями пользователя при оценке их точности и устойчивости;
- блок оценки качества **223**, предназначенный для оценки характеристики точности (по набору метрик, в частности могут использоваться метрики ГОСТ) и устойчивости моделей, разработанных и обученных пользователем, с использованием синтетических данных;
- блок выбора решения **224**, предназначенный для ранжирования различных реализации моделей, созданных пользователем, с точки зрения соответствия многомерным критериям качества.

Блок синтетических данных **221** выполнен с возможностью приёма-передачи данных из блока автоматической генерации структуры программных инструкций **210** и блока метрологического анализа **222**.

Блок метрологического анализа **222** выполнен с возможностью приёма-передачи данных из блока синтетических данных **211** и блока оценки качества **223**.

Блок оценки качества **223** выполнен с возможностью приёма-передачи данных из блока обработки сценариев **190** (в частности, блока трансляции **192**), блока метрологического анализа **222** и блока выбора решения **224**.

Блок выбора решения **224** выполнен с возможностью приёма-передачи данных из блока оценки качества **223**, блока планирования вычислений **130** и блока взаимодействия с пользователем **200**.

Описанное выше устройство может быть реализовано на текущем уровне развития техники и технологий за счёт использования к примеру элементов платформенных решений Национального центра когнитивных разработок (<https://actcognitive.org/platformy>): платформы работы с данными DataMall, платформы построения моделей SMILE, фреймворка автоматического машинного обучения FEDOT. В результате прототипирования обосновано, что предлагаемое устройство при помощи входящих в его состав блоков, обеспечивающих возможность приема/передачи данных с вышеперечисленными сервисами, при решении поставленных пользователями задач, способно работать с объемами данных не менее 100ТВ и обеспечивать и обрабатывать запрос на выполнение расчетов с использованием не менее чем 20480 ядер GPU одновременно для не менее чем 100 пользователей.

Описанное выше устройство обладает возможностью работы с большими данными, что является конкурентным преимуществом данного устройства. Большие данные – это данные, размещение и обработка которых на одном локальном компьютере физически невозможны. Однако устройство

работает с блоком, содержащим вычислительные ресурсы **110**, которые могут включать в себя и блоки распределённых вычислений **113**, в том числе, блоки, обеспечивающие поддержку работы с кластерами систем хранения больших данных на основе технологий Hadoop, Yarn, Spark, Storm, SAPHana, Cassandra, ClickHouse. Потому с технической точки зрения проблема больших данных снимается – устройство может так же эффективно работать с ними, как и с малыми – поскольку особенности организации их хранения учитываются при построении расписаний в блоке планирования вычислений **130**.

В качестве примера работы устройства создания модели машинного обучения на основе данных рассмотрим создание простой модели на данных:

1. осуществляют вход зарегистрированного пользователя через web-интерфейс блока взаимодействия с пользователем **200** и предоставляют ему доступ к функциям устройства в соответствии с установленными ему правами, определяемыми блоком управления пользователями **172**;
2. опционально загружают пользовательские данные посредством блока управления данными **171** или используют уже готовые данные, загруженные в устройство (при наличии соответствующих прав у пользователя), находящиеся в блоке биржи данных **142** посредством блока управления блоками бирж **150**;
3. с помощью блока взаимодействия с пользователем **200** получают от пользователя визуальную последовательность команд (Low-code-диаграмму) с описанием задачи машинного обучения, при необходимости выполняя настройку отдельных расчётных процедур посредством редактора исходного кода, включенного в блок взаимодействия с пользователем;
4. запускают процесс трансляции задачи, реализуемый блоком обработки сценариев **190**, который автоматически формирует набор программных

- инструкций для обучения моделей, удовлетворяющих условиям, полученным от пользователя, с помощью блока автоматической генерации структуры программных инструкций **210**, а также рассчитывают эффективный план его исполнения посредством блока планирования вычислений **130**;
5. программные инструкции для обучения моделей автоматически отправляют на исполнение в блок, содержащий вычислительные ресурсы **110** в соответствии с планом, выработанным блоком планирования вычислений **130** и контролируемым блоком управления ресурсами **120**;
  6. по завершении обучения моделей автоматически выполняют анализ их качества посредством блока интерпретации решений **220**;
  7. предоставляют пользователю в интерфейсе блока взаимодействия с пользователем **200** отчёт о работе устройства, включая ранжированный (в соответствии с критериями качества) перечень моделей, их параметров, оценок точности и времени расчётов;
  8. получают выбор пользователя подходящую ему модель и размещают ее в бирже моделей **141**, устанавливая права и условия ее использования другими пользователями. Таким образом, модель становится доступной как веб-сервис, к которому можно обращаться из других приложений посредством Интернет для решения прикладных задач.

В заключение следует отметить, что приведённые в описании сведения являются примерами, которые не ограничивают объём настоящего технического решения, определённого формулой.

## Формула

1. Устройство создания модели машинного обучения на основе данных, которое содержит:

а) блок взаимодействия с пользователем, выполненный с возможностью

обеспечения интерфейса человеко-компьютерного взаимодействия, содержащего средства для конструирования, разработки и исполнения композитных приложений, мониторинга вычислительных процессов и формирования отчетов на основе выполненных расчетов;

приема и передачи информации в блок управления окружением, блок обработки сценариев 190 и блок интерпретации решений;

б) блок управления окружением, выполненный с возможностью регистрации и управления правами пользователей; управления пользовательскими данными и сервисами, обеспечивающими пользователям доступ для работы с устройством;

приема и передачи информации в блок внутреннего взаимодействия и блок взаимодействия с пользователем;

в) блок бирж данных, содержащий информацию о биржах, предназначенных для каталогизации моделей, данных и задач, реализуемых пользователями системы и представленных в унифицированной форме, обеспечивающей их использование пользователями в соответствии с их правами, и выполненный с возможностью приема и передачи информации в блок управления блоками бирж;

г) блок управления блоками бирж, выполненный с возможностью обеспечения каталогизации и структурирования задач, данных и моделей, созданных пользователями и хранящихся в биржах блока хранения данных бирж,

приема и передачи информации в блок встраивания моделей и данных и блок внутренних взаимодействий;

д) блок обработки сценариев, выполненный с возможностью анализа формального описания задачи пользователя на создание модели машинного обучения, поступающей через блок взаимодействия с пользователем,

автоматического преобразование полученной задачи в исполнимую форму программных инструкций для обучения модели посредством блока автоматической генерации структуры программных инструкций, выполненный с возможностью приема и передачи информации в блок внутренних взаимодействий, блок взаимодействия с пользователем, блок планирования вычислений и блок интерпретации решений;

е) блок автоматической генерации структуры программных инструкций, соответствующего заданным пользователем критериям, выполненный с возможностью приема и передачи информации в блок обработки сценариев и блок интерпретации решения;

ж) блок планирования вычислений, выполненный с возможностью построения оптимального расписания исполнения всех элементов программных инструкций таким образом, чтобы обеспечить выполнение критериев пользователя при минимальной загрузке вычислительных ресурсов,

приема и передачи информации в блок управления ресурсами и блок обработки сценариев;

з) блок вычислений, содержащий вычислительные ресурсы, обеспечивающие распределённое выполнение расчётов в целях обучения модели, посредством запуска программных инструкций, сформированного блоком обработки сценариев в соответствии с расписанием блока планирования вычислений, выполненный с возможностью приема и передачи информации в блок управления ресурсами;

и) блок управления ресурсами, выполненный с возможностью обеспечения

загрузки и исполнения расчётных задач для обучения и использования моделей, в форме программных инструкций на вычислительных ресурсах,

приема и передачи информации в блок вычислений, содержащий вычислительные ресурсы и блок планирования вычислений;

к) блок встраивания моделей и данных, выполненный с возможностью

размещения в блоке управления блоками бирж новых данных и моделей, созданных вне данного устройства, выполненный с возможностью приема и передачи информации в блок управления блоками бирж;

л) блок интерпретации решений, выполненный с возможностью обеспечения контроля качества обученных моделей путём оценки их точности по сравнению с эталонными решениями, полученными с помощью блока автоматической генерации структуры программных инструкций на синтетических массивах данных,

приема и передачи информации в блок автоматической генерации структуры программных инструкций, блок обработки сценариев, блок планирования вычислений и блок взаимодействия с пользователем;

м) блок внутренних взаимодействий, выполненный с возможностью

организации совместной работы всех внутренних блоков и маршрутизации потоков данных в устройстве,

приема и передачи информации в блок управления блоками бирж, блок управления окружением и блок обработки сценариев;

2. Устройство по п.1, в котором блок, содержащий вычислительные ресурсы содержит:

- блок, содержащий физические вычислительные ресурсы, представляющие собой высокопроизводительные комплексы с собственной системой управления, предназначенные для реализации ресурсоёмких расчётов для обучения моделей, распараллеливаемые с учётом специфики их архитектуры, выполненный с возможностью приема и передачи данных/параметров в блок управления ресурсами;

- блок, связи с облачными вычислительными ресурсами, представляющими собой по меньшей мере публичные, частные облака, предназначенные для обеспечения использования обученных моделей как сервисов, а также для хранения промежуточных данных расчётов, выполненный с возможностью приема и передачи данных/параметров в блок управления ресурсами;

- блок распределённых вычислений, предназначенный для обеспечения доступа к распределённо хранящимся данным под единой системой управления на основе по меньшей мере пакетной, потоковой модели, выполненный с возможностью приема и передачи данных/параметров в блок управления ресурсами;

3. Устройство по п.1, в котором блок управления окружением содержит по меньшей мере:

- блок управления данными, предназначенный для обеспечения индексации данных в соответствии с заданной структурной моделью, а также для распределения по различным вычислительным ресурсам;

- блок управления пользователями, предназначенный для регистрирования в системе новых пользователей, установке прав пользователей, доступа к данным, задачам и ресурсам, а также для определения статистики их использования;

- блок управления сервисами, предназначенный для обеспечения доступа пользователя к данным, моделям и задачам как к цифровым сущностям, отображаемым в среде разработки, выполненный с

возможностью приема и передачи данных/параметров в блок взаимодействия с пользователем.

4. Устройство по п.1, в котором блок, включающий биржи, содержит по меньшей мере:

- биржу моделей, которая содержит модели машинного обучения, созданные разными пользователями, в унифицированном представлении, и делает доступными их другим пользователям в соответствии с их правами;

- биржу данных, которая содержит входные, выходные и синтетические данные, сформированные разными пользователями, в унифицированном представлении, и делает доступными их другим пользователям в соответствии с их правами;

- биржу задач, которая содержит описания прикладных задач машинного обучения, содержащих одну или несколько моделей, а также соответствующие им данные, при этом задачи формируются разными пользователями, хранятся в унифицированном представлении, и доступны другим пользователям в соответствии с их правами.

5. Устройство по п.1, в котором блок интерпретации решений содержит по меньшей мере:

- блок синтетических данных, предназначенный для обеспечения посредством вероятностных моделей и бутстрепирования генерации массивов данных, обладающих теми же свойствами, что и исходные, для формирования условий проверки качества работы обученных моделей, выполненный с возможностью приема и передачи данных/параметров в блок автоматической генерации структуры программных инструкций и блок метрологического анализа;

- блок метрологического анализа, реализованный на основе синтетических данных с помощью функций блока автоматической генерации структуры программных инструкций и предназначенный для постройки эталонных моделей, которые предназначены для

сопоставления с моделями пользователя при оценке их точности и устойчивости, выполненный с возможностью приема и передачи данных/параметров в блок синтетических данных и блок оценки качества;

- блок оценки качества, предназначенный для оценки характеристики точности по набору гостированных метрик и устойчивости моделей, разработанных и обученных пользователем, с использованием синтетических данных, выполненный с возможностью приема и передачи данных/параметров в блок обработки сценариев, блок метрологического анализа и блок выбора решения;

- блок выбора решения, предназначенный для ранжирования различных реализации моделей, созданных пользователем, с точки зрения соответствия многомерным критериям качества, выполненный с возможностью приема и передачи данных/параметров в блок оценки качества, блок планирования вычислений и блок взаимодействия с пользователем.

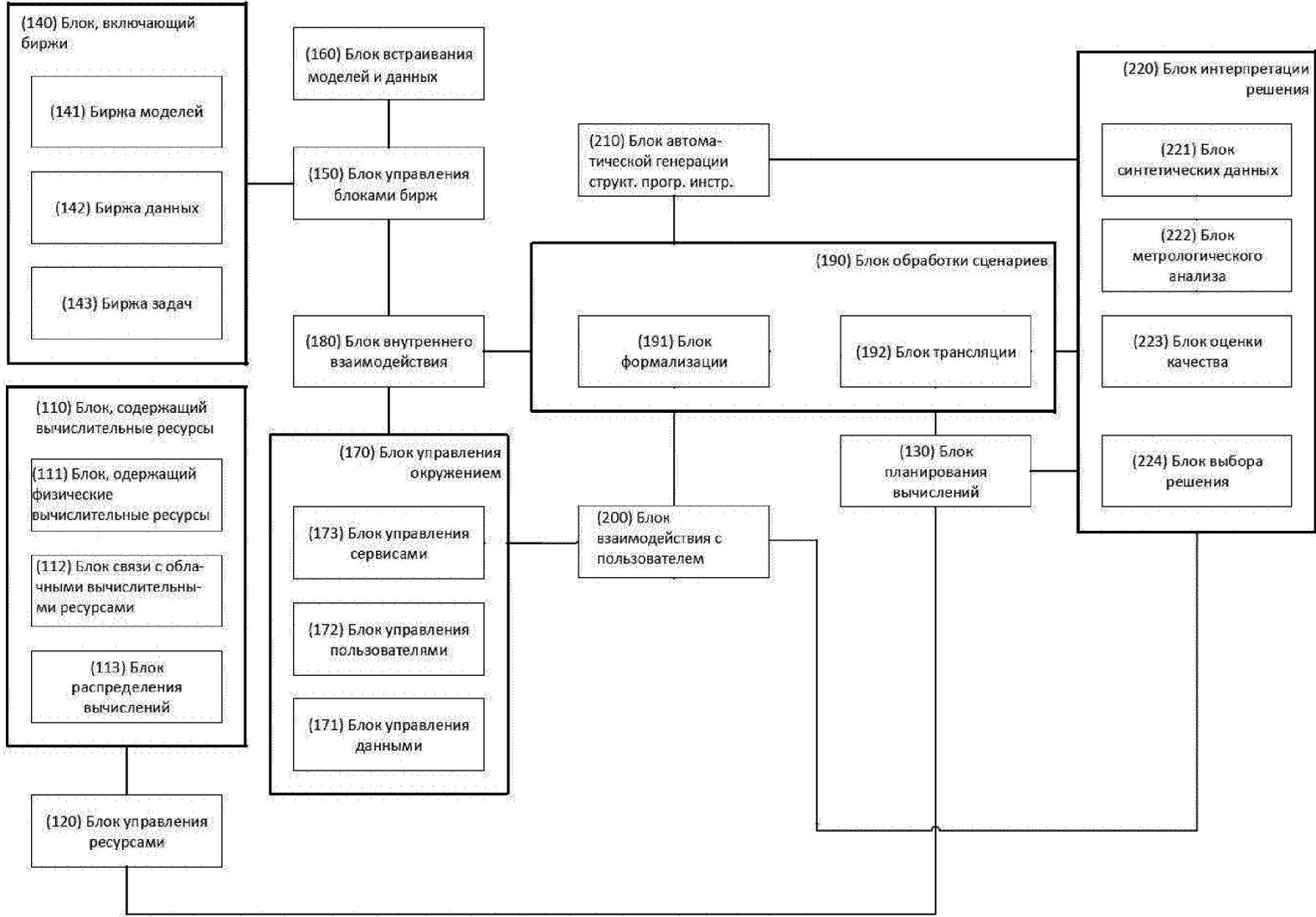
6. Устройство по п.1, в котором блок обработки сценариев содержит по меньшей мере:

- блок формализации, предназначенный для обеспечения высокоуровневой формализации и выполнения задачи пользователя по подготовке входных и выходных данных, формированию структуры модели в виде композитного приложения и запуску модели на обучение, выполненный с возможностью приема и передачи данных/параметров в блок внутреннего взаимодействия, блок трансляции, блок автоматической генерации структуры программных инструкций и блок взаимодействия с пользователем;

- блок трансляции, предназначенный для обеспечения трансляции, где трансляция представляет собой перевод высокоуровневого представления композитного приложения задачи в виде графа в детализированную исполнимую форму с определением реализаций всех расчётных функций, выполненный с возможностью приема и передачи данных/параметров в блок

формализации, блок планирования вычислений и блок интерпретации решения.

Фиг. 1.



**ОТЧЕТ О ПАТЕНТНОМ ПОИСКЕ**

(статья 15(3) ЕАПК и правило 42 Патентной инструкции к ЕАПК)

Номер евразийской заявки:

**202392893****А. КЛАССИФИКАЦИЯ ПРЕДМЕТА ИЗОБРЕТЕНИЯ:**

МПК:

**G06N 3/08** (2023.01)  
**G06N 7/04** (2006.01)  
**G06N 20/00** (2019.01)  
**G05B 13/04** (2006.01)

СПК:

**G06N 3/08**  
**G06N 7/04**  
**G06N 20/00**  
**G05B 13/04**

**Б. ОБЛАСТЬ ПОИСКА:**

G06N 3/00-3/08, 7/00-7/04, 20/00, G05B 13/00-13/04

Электронная база данных, использовавшаяся при поиске (название базы и, если возможно, используемые поисковые термины)  
 Google Patents, Espacenet, (ИС «Поисковая платформа» Роспатент), ЕАПАТИС

**В. ДОКУМЕНТЫ, СЧИТАЮЩИЕСЯ РЕЛЕВАНТНЫМИ**

| Категория* | Ссылки на документы с указанием, где это возможно, релевантных частей   | Относится к пункту № |
|------------|---|----------------------|
| A          | US 20190080347 A1 (ADOBE INC.) 2019-03-14   | 1-6                  |
| A          | US 20190121889 A1 (PURE STORAGE, INC.) 2019-04-25   | 1-6                  |
| A          | RU 2411574 C2 (БУХАНОВСКИЙ АЛЕКСАНДР ВАЛЕРЬЕВИЧ, ВАСИЛЬЕВ ВЛАДИМИР НИКОЛАЕВИЧ, НЕЧАЕВ ЮРИЙ ИВАНОВИЧ) 2011-02-10   | 1-6                  |
| A          | RU 2596992 C2 (ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ "САНКТ-ПЕТЕРБУРГСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ" (УНИВЕРСИТЕТ ИТМО)) 2016-09-10   | 1-6                  |
| A          | RU 2014150572 A (ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ "САНКТ-ПЕТЕРБУРГСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ" (УНИВЕРСИТЕТ ИТМО)) 2016-07-10 | 1-6                  |

 последующие документы указаны в продолжении графы

\* Особые категории ссылочных документов:

«А» - документ, определяющий общий уровень техники

«D» - документ, приведенный в евразийской заявке

«E» - более ранний документ, но опубликованный на дату подачи евразийской заявки или после нее

«O» - документ, относящийся к устному раскрытию, экспонированию и т.д.

"P" - документ, опубликованный до даты подачи евразийской заявки, но после даты испрашиваемого приоритета"

«Т» - более поздний документ, опубликованный после даты приоритета и приведенный для понимания изобретения

«X» - документ, имеющий наиболее близкое отношение к предмету поиска, порочащий новизну или изобретательский уровень, взятый в отдельности

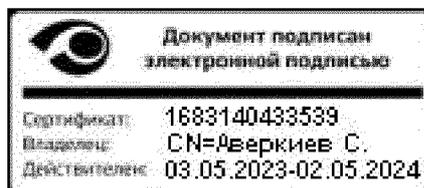
«Y» - документ, имеющий наиболее близкое отношение к предмету поиска, порочащий изобретательский уровень в сочетании с другими документами той же категории

«&amp;» - документ, являющийся патентом-аналогом

«L» - документ, приведенный в других целях

Дата проведения патентного поиска: 14 февраля 2024 (14.02.2024)

Уполномоченное лицо:  
 Начальник Управления экспертизы



С.Е. Аверкиев