ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

ПРЕДВАРИТЕЛЬНЫЙ НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ПНСТ 845— 2023

Искусственный интеллект

ТЕХНИЧЕСКАЯ СТРУКТУРА ФЕДЕРАТИВНОЙ СИСТЕМЫ МАШИННОГО ОБУЧЕНИЯ

Издание официальное

Москва
Российский институт стандартизации
2023

Предисловие

1 РАЗРАБОТАН Научно-образовательным центром компетенций в области цифровой экономики Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова» (МГУ имени М.В. Ломоносова) и Обществом с ограниченной ответственностью «Институт развития информационного общества» (ИРИО)

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 164 «Искусственный интеллект»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 13 декабря 2023 г. № 92-пнст

4 Настоящий стандарт разработан с учетом основных нормативных положений международного стандарта МСЭ F.748.13 (2021) «Техническая структура федеративной системы МО» (Recommendation ITU-T F.748.13 (2021), Technical framework for a shared machine learning system, NEQ)

Правила применения настоящего стандарта и проведения его мониторинга установлены в ГОСТ Р 1.16—2011 (разделы 5 и 6).

Федеральное агентство по техническому регулированию и метрологии собирает сведения о практическом применении настоящего стандарта. Данные сведения, а также замечания и предложения по содержанию стандарта можно направить не позднее чем за 4 мес до истечения срока его действия разработчику настоящего стандарта по адресу: 119991 Москва, Ленинские горы, д. 1 и в Федеральное агентство по техническому регулированию и метрологии по адресу: 123112 Москва, Пресненская набережная, д. 10, стр. 2.

В случае отмены настоящего стандарта соответствующая информация будет опубликована в ежемесячном информационном указателе «Национальные стандарты» и также будет размещена на официальном сайте Федерального агентства по техническому регулированию и метрологии (www.rst.gov.ru)

Содержание

Введение

Настоящий стандарт определяет роли, устанавливает технические требования и требования по безопасности для федеративной системы машинного обучения, а также описывает технические архитектуры, функциональные компоненты и процедуры обработки федеративной системы машинного обучения при централизованном и децентрализованном режимах работы.

В настоящий стандарт включены дополнительные по отношению к МСЭ F.748.13 определения и положения, приведенные в [1]. Это позволяет гармонизировать настоящий стандарт с принятыми ранее национальными стандартами и предварительными национальными стандартами в области искусственного интеллекта.

Доступ к МСЭ F.748.13 можно получить по адресу: http://handle.itu.int/ (уникальный идентификатор — http://handle.itu.int/11.1002/1000/14682).

Все рекомендации МСЭ и другие источники могут быть пересмотрены, поэтому пользователям настоящего стандарта предлагается изучить возможность применения последнего издания рекомендаций и других справочных документов (перечень действующих в настоящее время рекомендаций МСЭ регулярно публикуется). Ссылка на рекомендацию МСЭ в рамках настоящего стандарта не придает ему как отдельному документу статус стандарта.

ПРЕДВАРИТЕЛЬНЫЙ НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Искусственный интеллект

ТЕХНИЧЕСКАЯ СТРУКТУРА ФЕДЕРАТИВНОЙ СИСТЕМЫ МАШИННОГО ОБУЧЕНИЯ

Artificial intelligence. Technical framework for a shared machine learning system

Срок действия — с 2024—02—01
до 2027—02—01

1 Область применения

В настоящем стандарте установлено понятие федеративной системы машинного обучения, а также определены роли, технические требования и требования по безопасности, технические архитектуры, функциональные компоненты и процедуры обработки федеративной системы машинного обучения при централизованном и децентрализованном режимах работы. В приложении А также приведено описание вариантов использования федеративных систем машинного обучения.

2 Термины и определения

В настоящем стандарте применены следующие термины с соответствующими определениями:

2.1 **машинное обучение** (machine learning): Процесс оптимизации параметров модели с помощью вычислительных методов таким образом, чтобы поведение модели отражало данные и/или опыт.

Примечание — См. [1], пункт 3.3.5.

2.2 **доверенная среда исполнения** (trusted execution environment): Защищенная область процессора, которая обеспечивает хранение, обработку и защиту данных и целостность исполняемого кода в изолированной и надежной среде.

Примечание — Такая среда поддерживает изолированное защищенное выполнение авторизованного программного обеспечения системы безопасности, которое позволяет обеспечить сквозную безопасность посредством исполнения в защищенном режиме аутентифицированного кода и обеспечения конфиденциальности, аутентичности, защиты персональных данных, целостности системы и прав доступа к данным.

Примечание — См. [2], пункт 3.2.9.

2.3 **федеративное (совместное) машинное обучение** (shared machine learning): Парадигма машинного обучения, позволяющая агрегировать принадлежащие ряду сторон данные и обеспечивать многостороннюю защиту персональных данных в тех ситуациях, когда различные поставщики данных и вычислительная платформа не доверяют друг другу.

2.4 **безопасные многосторонние вычисления** (multi-party computation): Подраздел криптографии, занимающийся созданием методов, позволяющих сторонам совместно вычислять значение функции на основе индивидуально хранимых сторонами частей входных данных, сохраняя при этом конфиденциальность этих частей входных данных.

2.5 **удаленная аттестация** (remote attestation): Метод, с помощью которого вычислительный узел (клиент) осуществляет аутентификацию конфигурации своего оборудования и программного обеспечения для удаленного вычислительного узла (сервера).

Примечание — Цель удаленной аттестации — дать возможность одной удаленной системе (запрашивающей доказательства) определить уровень доверия к целостности платформы другой системы (заявляющей о такой целостности).

Издание официальное

3 Сокращения

В настоящем стандарте применены следующие сокращения:

MO — машинное обучение;

API — интерфейс программирования приложений (application programming interface);

QPS — количество запросов в секунду (queries per second);

SML — федеративное машинное обучение (shared machine learning);

SMS — служба коротких сообщений (short messaging service);

TEE — доверенная среда исполнения (trusted execution environment);

WOE — весомость доказательств (weight of evidence).

4 Соглашения по терминологии

В настоящем стандарте:

- ключевые слова «требуется, чтобы» означают требование, которое должно строго соблюдаться и отклонение от которого не допускается, если будет сделано заявление о соответствии настоящему стандарту;

- ключевое слово «рекомендуется» означает требование, которое рекомендуется, но не является абсолютно необходимым. Таким образом это требование не является обязательным для заявления о соответствии настоящему стандарту;

- ключевые слова «может опционально» означают необязательное требование, которое является допустимым, но при этом не подразумевается, что оно в каком-либо смысле рекомендуется. Данная формулировка не подразумевает ни обязанности предлагаемого поставщиком варианта реализации предоставить такую опцию, ни возможности опционального подключения такой функциональной возможности оператором сети/поставщиком услуг. Она означает то, что поставщик может опционально предоставить эту функциональную возможность и по-прежнему заявлять о соответствии спецификации.

5 Общее представление о федеративной системе машинного обучения

В федеративных системах МО несколько участников совместно используют зашифрованные данные и/или обмениваются параметрами моделей с целью обеспечить безопасность данных и защиту персональных данных. Для обеспечения наиболее эффективного использования данных зашифрованные данные каждой стороны и/или предоставленные ею параметры моделей собираются и используются для обучения модели федеративного МО. Модели федеративного МО продолжают обучаться для самооптимизации, а участники или иные лица, авторизованные на доступ к модели, могут вводить информацию для получения результатов или прогнозов на основе совместно используемых значений. Федеративные системы МО могут, например, применяться (не ограничиваясь ими) в мультимедийных и игровых приложениях.

Федеративное МО может использоваться в централизованном и децентрализованном режимах. Централизованный режим — это решение для многостороннего совместного использования зашифрованных данных и для обучения на результатах слияния данных в доверенной среде исполнения. Децентрализованный режим — это решение для совместного использования и обучения несколькими участниками, основанное на безопасных многосторонних вычислениях, при котором осуществляется обмен неоригинальными данными, не раскрывающими персональные данные.

При использовании централизованного режима сбор данных и обучение на них осуществляет доверенная третья сторона. Данный режим можно применять в тех ситуациях, когда участники готовы обмениваться данными, не содержащими персональные данные, а стоимость доступа к услугам низкая. Его также можно применять в тех случаях, когда требуется проведение сложных вычислений. В централизованном режиме могут также поддерживаться любые алгоритмы, развертывание кластера и централизованные вычисления.

При использовании децентрализованного режима применяют методы безопасных многосторонних вычислений для обмена данными или параметрами моделей с целью выполнения задач обучения. Данный режим можно применять в тех ситуациях, в которых участникам требуется строгая защита персональных данных или когда у участников имеются большие объемы локальных данных, так как в

рамках обучения модели МО стороны обмениваются параметрами моделей, а не исходными данными. Безопасные многосторонние вычисления подходят в случае простых вычислений, так как они способны поддерживать меньше алгоритмов по сравнению с решением на основе доверенной среды исполнения.

6 Роли в федеративной системе машинного обучения

6.1 Обзор

Роли в федеративной системе МО и взаимодействие между ними показаны на рисунке 1.

В федеративной системе МО участники могут выполнять ряд ролей, включая роли инициатора задачи, поставщика данных, вычислительной платформы и получателя результатов. Следует иметь в виду, что, хотя на рисунке 1 показаны только два поставщика данных, однако при реальном развертывании решения таких поставщиков данных может быть много. Одна и та же сторона может выступать в качестве исполнителя нескольких ролей, например: один из поставщиков данных может быть инициатором задачи и/или получателем результата.

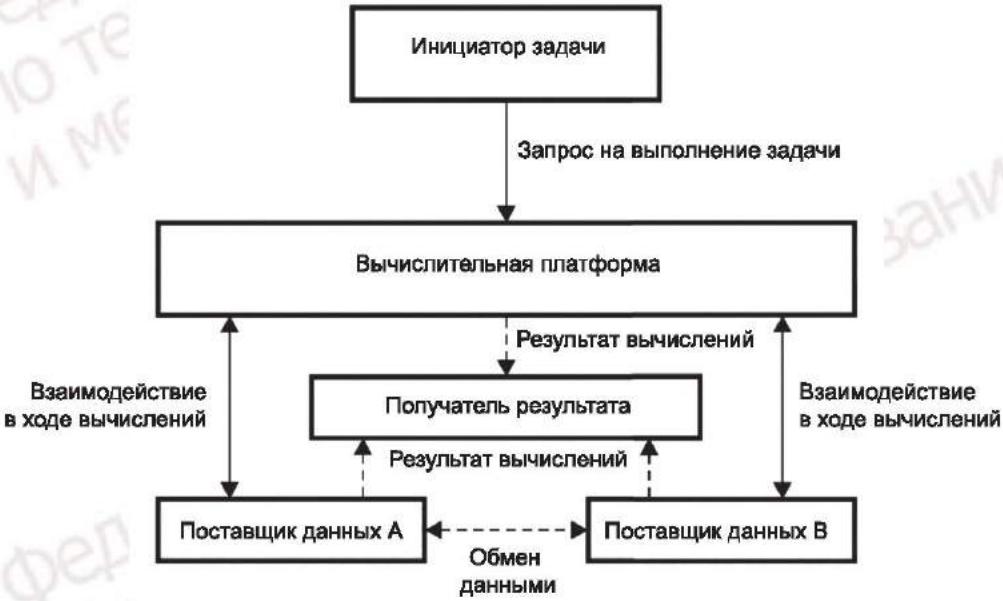

Рисунок 1 — Роли в федеративной системе МО

6.2 Поставщик данных

Поставщик данных является обладателем данных и предоставляет их в качестве входных данных вычислительной платформе или другому поставщику данных. Следует обратить внимание на то, что поставщики данных также располагают вычислительными ресурсами.

В централизованном режиме данные шифруются и передаются от поставщика данных вычислительной платформе.

В децентрализованном режиме проводится обработка конфиденциальных сведений и/или параметров модели, обмен которыми осуществляется между поставщиками данных при координации со стороны вычислительной платформы.

И в том, и в другом режиме обмениваемые данные представляют собой полученные в результате расчетов коэффициенты, а не сами исходные данные. Ни у платформы вычислений, ни у поставщиков данных не должно быть возможности получить какую-либо информацию об исходных данных из этих коэффициентов.

6.3 Вычислительная платформа

Вычислительная платформа получает от инициатора задачи запрос на выполнение задачи и отправляет зашифрованный результат получателю. Она взаимодействует с поставщиками данных с целью выполнения вычислительных задач.

В централизованном режиме вычислительная платформа получает зашифрованные данные от поставщиков данных, расшифровывает их и выполняет вычисления в доверенной среде исполнения.

В децентрализованном режиме вычислительная платформа распределяет выполнение вычислительной задачи между участниками и координирует деятельность поставщиков данных по обмену параметрами и/или конфиденциальными сведениями между собой и по индивидуальному выполнению вычислительных задач.

Вычислительная платформа объединяет вычислительные алгоритмы.

6.4 Получатель результата

В централизованном режиме по завершении вычислительной задачи вычислительная платформа посылает зашифрованный результат получателю, который может его расшифровать и получить окончательный результат.

В децентрализованном режиме поставщики данных посылают результаты вычислений получателю, который их объединяет и получает окончательный результат.

Роль получателя результатов может выполняться и поставщиками данных, и инициатором задачи, и вычислительной платформой.

6.5 Инициатор задачи

Инициатор задачи инициирует вычислительную задачу на вычислительной платформе.

Роль инициатора задачи может выполняться вычислительной платформой или поставщиками данных.

7 Технические требования к федеративной системе машинного обучения

7.1 Основные функциональные требования

7.1.1 Функциональные требования к управлению данными

Функциональные требования к управлению данными для федеративной системы МО включают следующее:

- федеративная система МО должна поддерживать управление метаданными и их отображение;
- федеративная система МО должна поддерживать функцию авторизации использования данных;
- федеративная система МО должна поддерживать функцию синхронизации данных (data alignment);
- рекомендуется, чтобы федеративная система МО поддерживала аудит поведения при использовании данных;
- рекомендуется, чтобы федеративная система МО поддерживала хранение данных в зашифрованном виде.

7.1.2 Функциональные требования к управлению алгоритмами

Функциональные требования по управлению алгоритмами для федеративной системы МО включают следующее:

- федеративная система МО должна поддерживать по крайней мере один широко распространенный алгоритм МО и прогнозирования, для которого есть возможность обеспечить защиту персональных данных, например: линейную регрессию, логистическую регрессию, древовидную модель, глубокую нейронную сеть, графовую нейронную сеть;
- федеративная система МО должна поддерживать функции настройки параметров для получения более качественных результатов обучения;
- федеративная система МО должна поддерживать по крайней мере один метод секционирования данных (горизонтальное секционирование или вертикальное секционирование); рекомендуется, чтобы она поддерживала оба метода секционирования данных одновременно;
- федеративная система МО должна поддерживать модели, предусматривающие защиту от так называемых получестных атак.

Примечание — Получестная атака (semi-honest attack) — это атака, при которой одна сторона, следуя установленному протоколу, пытается извлечь больше информации из сообщений, получаемых от других сторон, чем это допустимо;

- рекомендуется, чтобы федеративная система МО поддерживала учитывающую потребности защиты персональных данных функцию, обеспечивающую сведение данных при повторном объединении вертикально секционированных данных;

- рекомендуется, чтобы федеративная система МО поддерживала безопасные функции анализа данных, такие как статистический анализ максимальных и минимальных значений, средней дисперсии;

- рекомендуется, чтобы федеративная система МО поддерживала функции предварительной обработки данных и оценки эффективности алгоритмов МО, включая разделение набора данных, заполнение пропущенных значений, метод весомости доказательств, оценку корреляции признаков и другие инструменты для повышения показателей производительности моделей МО.

7.1.3 Функциональные требования к управлению вычислениями

Функциональные требования к управлению вычислениями для федеративной системы МО включают следующее:

- федеративная система МО должна предоставлять базовые возможности для управления задачами, такие как создание и отмена задач;

- рекомендуется, чтобы федеративная система МО предоставляла развитые возможности для управления задачами, такими как мониторинг хода выполнения задач, постановка задач в очередь, анализ исторической информации о задачах;

- рекомендуется, чтобы федеративная система МО поддерживала возможности для управления распределенными ресурсами и планирования задач.

7.2 Требования к масштабируемости

Требования к масштабируемости для федеративной системы МО включают следующее:

- рекомендуется, чтобы федеративная система МО обладала хорошей масштабируемостью и обеспечивала добавление новых функциональных компонентов в соответствии с деловыми потребностями;

- рекомендуется, чтобы федеративная система МО поддерживала возможность доступа пользователей к системе через интерфейс прикладного программирования.

7.3 Требования к надежности

Требования к надежности для федеративной системы МО включают следующее:

- федеративная система МО должна гарантировать доступность системы и предотвращать отключения из-за некорректных входных данных;

- федеративная система МО должна быть способна выполнять автоматическое аварийное восстановление после сбоев (таких как сбой сервера, сбой жесткого диска, сбой сети, выключение, перезапуск), включая резервное копирование и восстановление данных и т. д.;

- рекомендуется, чтобы федеративная система МО поддерживала кластеризацию обучения и прогнозирования, а также поддерживала проведение аварийного восстановления из нескольких компьютерных залов, с целью обеспечения катастрофоустойчивости и возможности переключения сервиса на другой ресурс в случае отказа, тем самым повышая доступность системы;

- рекомендуется, чтобы федеративная система МО для повышения надежности системы поддерживала такие эксплуатационные возможности и возможности для технического обслуживания, как предварительное тестирование и развертывание по модели «канареечного релиза» (canary or grayscale release), мониторинг и оповещение, а также отслеживание ссылок.

7.4 Требования к совместимости

Требования к совместимости для федеративной системы МО включают следующее:

- федеративная система МО должна обеспечивать использование других алгоритмов МО;

- рекомендуется, чтобы федеративная система МО поддерживала обратную совместимость в процессе обновления, включая совместимость между системными модулями и совместимость между системами и файлами;

- рекомендуется, чтобы федеративная система МО поддерживала требования к развертыванию в различных средах, таких как облачная среда, виртуальная машина, физическая машина;

- рекомендуется, чтобы федеративная система МО поддерживала нормальную работу решения на основе доверенной среды исполнения в различных широко распространенных доверенных средах исполнения.

7.5 Требования к производительности

Требования к производительности для федеративной системы МО включают следующее:

- рекомендуется, чтобы федеративная система МО поддерживала распределенное обучение для того, чтобы увеличить объем данных, используемых в процессе обучения, и сократить время обучения;
- рекомендуется, чтобы федеративная система МО была способна повысить производительность, измеряемую в запросах в секунду, за счет горизонтального масштабирования.

7.6 Требования к удобству использования

Требования к удобству использования для федеративной системы МО включают следующее:

- федеративная система МО должна предоставлять полные инструкции по развертыванию и эксплуатации, чтобы облегчить пользователям понимание системы, доступ к ней и ее использование;
- рекомендуется, чтобы федеративная система МО предоставляла платформенный интерфейс для снижения стоимости обучения пользователей и использования системы;
- рекомендуется, чтобы федеративная система МО предоставляла платформу разработки алгоритмов для того, чтобы на ее основе пользователи могли разрабатывать собственные алгоритмы, отвечающие ограничениям по безопасности;
- рекомендуется, чтобы федеративная система МО была спроектирована таким образом, чтобы сократить продолжительность перерывов в работе пользователей во время выполнения обновлений системы.

8 Требования к безопасности для федеративной системы машинного обучения

8.1 Требования к аутентификации

Требования к аутентификации для федеративной системы МО включают следующее:

- федеративная система МО должна поддерживать функции аутентификации личности пользователей, которые получают доступ к федеративной системе МО. К пользователям относятся поставщики данных, системные пользователи и получатели результатов;
- федеративная система МО должна поддерживать аутентификацию сертификатов для поставщиков данных и получателей результатов;
- рекомендуется, чтобы федеративная система МО поддерживала две или более комбинации технологий (таких как проверка пароля, проверка с помощью электронной почты, проверка с помощью SMS и т. д.) при аутентификации личности пользователей системы.

8.2 Требования к управлению доступом

Требования к управлению доступом для федеративной системы МО включают следующее:

- федеративная система МО должна поддерживать механизм авторизации для пользователей, чтобы гарантировать авторизованность использования их данных в федеративном машинном обучении;
- рекомендуется, чтобы федеративная система МО в полной мере поддерживала роли и системы авторизации. Федеративная система МО должна явным образом назначать роли и устанавливать права доступа после входа пользователя в систему;
- рекомендуется, чтобы федеративная система МО требовала от пользователя в соответствии с политикой МО повторной аутентификации или повторной активации сеанса работы после продолжительного бездействия.

8.3 Требования к аудиту безопасности

Требования к аудиту безопасности для федеративной системы МО включают следующее:

- федеративная система МО должна поддерживать функции протоколирования и анализа журналов аудита для операций с данными, выполняемых при обращении пользователей к федеративной системе МО;
- рекомендуется, чтобы федеративная система МО поддерживала функцию отслеживания и аудита истории операций для основных операций в системе.

8.4 Требования к безопасности данных

Требования к безопасности данных для федеративной системы МО включают следующее:

- федеративная система МО должна поддерживать функцию передачи и хранения важных данных в зашифрованном виде;
- обрабатываемые данные не должны содержать чувствительные (в т. ч. секретные и конфиденциальные) данные, обработка которых запрещена действующим законодательством;
- федеративная система МО должна поддерживать возможность ограничения передачи данных;
- федеративная система МО должна поддерживать защищенные протоколы передачи и/или защищенные каналы передачи, чтобы обеспечить безопасность и надежность каналов передачи данных и предотвратить возможные атаки;
- федеративная система МО должна обеспечивать свойства конфиденциальности, целостности и доступности данных и выборок данных, а также предотвращать незаконное получение данных неавторизованными пользователями;
- федеративная система МО должна уничтожать доверенную среду исполнения после завершения выполнения вычислительных задач в централизованном режиме.

8.5 Требования к защите персональных данных

Требования к защите персональных данных для федеративной системы МО включают следующее:

- федеративная система МО должна обеспечивать, чтобы имеющиеся у пользователей персональные данные не попадали в результате утечек к другим поставщикам данных, координаторам или пользователям;
- федеративная система МО должна обеспечивать наличие механизма, предотвращающего возможность каким-либо образом осуществлять идентификацию физического лица (с той или иной степенью точности) на основе информации, которой обмениваются поставщики данных;
- федеративная система МО должна предоставлять поставщикам данных возможность использования адекватных методов шифрования, обеспечивающих, что другие стороны, помимо получателя данных, не смогут извлечь персональные данные пользователя из пересылаемых зашифрованных данных, чтобы предотвратить такие возможные атаки, как взлом методом грубой силы, логические атаки и т. д.;
- в рамках безопасного многостороннего вычислительного решения федеративная система МО должна обеспечить, что имеющиеся у пользователя персональные данные остаются в локальном хранилище, а различные поставщики данных обмениваются только случайными числами и/или зашифрованными параметрами.

9 Техническая архитектура, функциональные компоненты и процедура обработки федеративной системы машинного обучения в централизованном режиме

9.1 Техническая архитектура федеративной системы МО в централизованном режиме

В централизованном режиме работы федеративной системы МО локальные данные ряда участников обрабатываются, шифруются и передаются в доверенную среду исполнения для обучения федеративной модели. Поскольку данные каждой стороны зашифрованы, их не может увидеть ни платформа, ни другие участники. Когда участники или иные лица имеют авторизацию на доступ к федеративной модели, они могут получить к ней доступ посредством интерфейсов прикладного программирования для ввода исходных данных и получения выходных данных или прогнозов на основе совместно используемых значений.

Перед предоставлением данных участник (поставщик данных) проверяет корректность и правильность работающего в доверенной среде исполнения программного обеспечения посредством удаленной аттестации. После этого участник и доверенная среда исполнения согласовывают механизм шифрования. Затем участник шифрует данные, используя согласованный механизм, и отправляет зашифрованные данные в доверенную среду исполнения. Для шифрования могут быть применены механизмы как симметричного, так и асимметричного шифрования.

Чтобы исключить утечку вычисленной модели, обученная модель должна быть зашифрована после завершения всех вычислений.

Техническая архитектура представлена на рисунке 2.

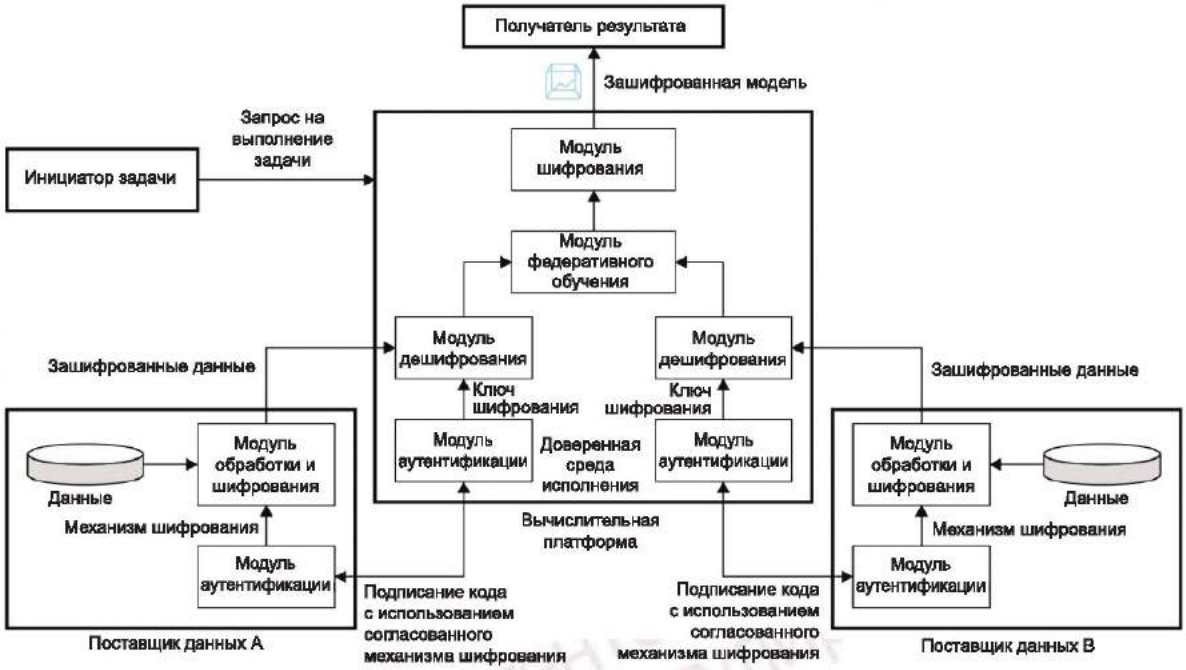

Рисунок 2 — Техническая архитектура федеративной системы МО в централизованном режиме

Техническая архитектура федеративной системы МО в централизованном режиме в основном состоит из вычислительной платформы и нескольких поставщиков данных. Вычислительная платформа включает модуль аутентификации, модули шифрования и дешифрования, а также модуль федеративного обучения. У каждого поставщика данных имеются модуль обработки и шифрования, модуль аутентификации и данные. В технической архитектуре модули являются функциональными компонентами.

Данные каждого поставщика данных обрабатываются и шифруются с помощью ключа шифрования, полученного от модуля аутентификации, а затем зашифрованные данные загружаются в доверенную среду исполнения вычислительной платформы. Модуль аутентификации вычислительной платформы расшифровывает зашифрованные данные с помощью ключа дешифрования, после чего отправляет расшифрованные данные в модуль федеративного обучения. Модуль федеративного обучения выполняет операцию федеративного МО с использованием расшифрованных данных, полученных от нескольких поставщиков данных.

9.2 Функциональные компоненты федеративной системы МО в централизованном режиме

9.2.1 Модуль аутентификации

Данный модуль отвечает за согласование механизма шифрования с каждым поставщиком данных и за предоставление ключа дешифрования модулю дешифрования. Ключ дешифрования используется для расшифровки зашифрованных данных, загруженных на платформу поставщиком данных. Модуль аутентификации вычислительной платформы отвечает за подписание программного кода, работающего в доверенной среде исполнения, и поддерживает проведение проверки программного кода поставщиком данных.

9.2.2 Модуль дешифрования

Данный модуль отвечает за расшифровку загруженных поставщиком данных зашифрованных данных с использованием ключа дешифрования. Ключ дешифрования может быть как асимметричным, так и симметричным.

9.2.3 Модуль федеративного обучения

Данный модуль отвечает за обучение модели на основе расшифрованных данных, полученных от нескольких поставщиков данных, с целью вычисления федеративной модели.

9.2.4 Модуль шифрования

Данный модуль отвечает за шифрование вычисленной модели после завершения всех вычислений.

Основу архитектуры поставщика данных составляют данные, модуль обработки и шифрования, а также модуль аутентификации. Имеется как минимум два поставщика данных.

9.2.5 Данные

Данные, предоставленные поставщиком данных, могут быть обработаны, зашифрованы и загружены на вычислительную платформу. Зашифрованы могут быть первичные данные, рассчитанные коэффициенты и любыми иные данные, которые необходимо защитить при передаче между участниками и вычислительной платформой.

9.2.6 Модуль обработки и шифрования

Данный модуль используется для обработки и шифрования данных на основе механизма шифрования, согласованного с вычислительной платформой, после чего зашифрованные данные загружаются на вычислительную платформу.

9.2.7 Модуль аутентификации

Модуль аутентификации поставщика данных используется для выполнения удаленной аутентификации поставщика данных и вычислительной платформы, включая согласование с вычислительной платформой механизма шифрования и проверку подписи кода.

Инициатор задачи отправляет на вычислительную платформу запрос на выполнение вычислительной задачи, и инициирует вычислительную задачу.

Получатель результата получает от вычислительной платформы зашифрованный результат и, расшифровав его, получает окончательный результат.

9.3 Процедуры обработки федеративной системы МО в централизованном режиме

В централизованном режиме как любые поставщики данных, так и вычислительная платформа могут инициализировать вычислительные задачи, после чего вычислительная платформа создает доверенную среду исполнения. Локальные данные, предоставляемые каждым поставщиком данных, могут быть обработаны, зашифрованы и загружены на вычислительную платформу. Платформа расшифровывает полученные от поставщиков данных зашифрованные данные в доверенной среде исполнения и выполняет обучение модели на основе расшифрованных данных для получения федеративной модели. Обработка данных, шифрование, дешифрование и шаги обучения могут повторяться несколько раз. По завершении вычислительной задачи доверенная среда исполнения уничтожается для обеспечения безопасности и защиты персональных данных.

Рекомендуемые процедуры обработки представлены на рисунке 3.

Технический процесс обработки в функционирующей в централизованном режиме федеративной системе МО включает следующие шаги:

- шаг 1: поставщики данных скачивают программные инструменты с вычислительной платформы и развертывают их;
- шаг 2: поставщики данных выполняют подготовку данных, включая их шифрование и авторизацию;
- шаг 3: поставщики данных загружают зашифрованные данные на вычислительную платформу;
- шаг 4: инициатор задачи инициирует вычислительные задачи на вычислительной платформе, включая обучаемую модель и алгоритмы;
- шаг 5: вычислительная платформа создает доверенную среду исполнения;
- шаг 6: вычислительная платформа расшифровывает зашифрованные данные в доверенной среде исполнения;
- шаг 7: вычислительная платформа в доверенной среде исполнения выполняет вычисления с использованием расшифрованных данных для получения результата вычислений;
- шаг 8: вычислительная платформа доставляет результат вычисления получателю результата;
- шаг 9: вычислительная платформа уничтожает доверенную среду исполнения и данные в ней.

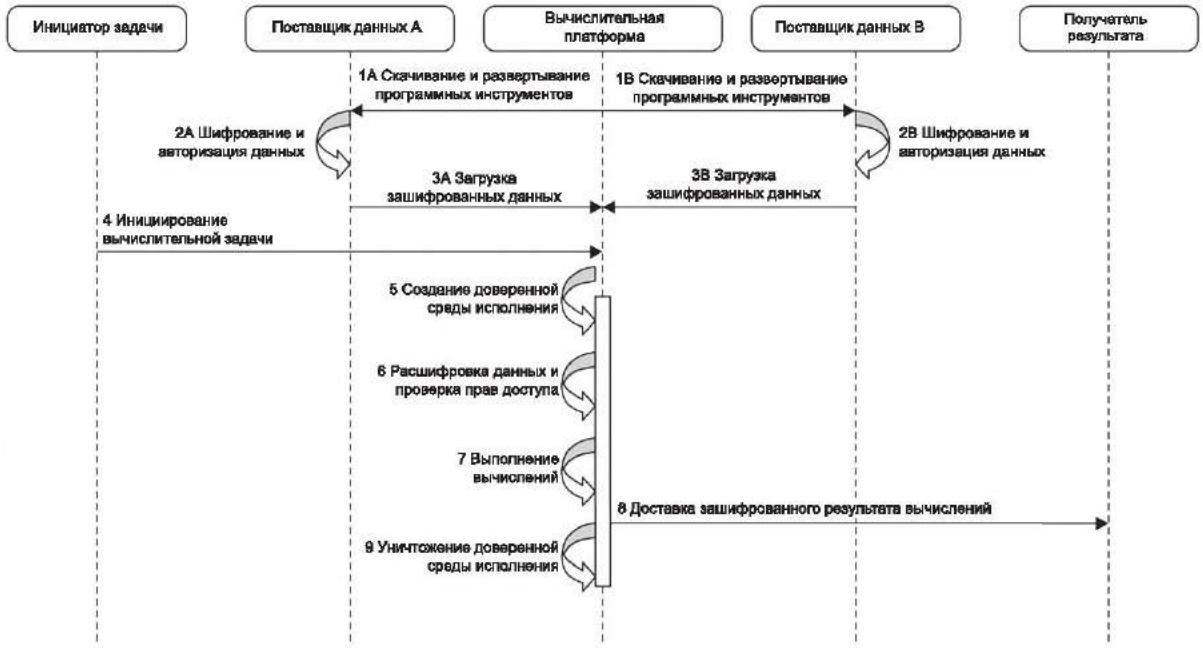

Рисунок 3 — Процедуры обработки федеративной системы МО в централизованном режиме

Следует иметь в виду, что описанные выше процедуры обработки являются типовыми. В условиях каждого конкретного развертывания решения порядок и детали каждой процедуры могут отличаться. Например, инициатор задачи может инициировать вычисление на первом шаге. Поставщики данных могут многократно загружать зашифрованные данные на вычислительную платформу. Относящиеся к обучению модели данные и обученная модель также могут быть «запечатаны» и сохранены вне вычислительной платформы для последующего повторного использования.

10 Техническая архитектура, функциональные компоненты и процедура обработки федеративной системы машинного обучения в децентрализованном режиме

10.1 Техническая архитектура федеративной системы МО в децентрализованном режиме

В децентрализованном режиме работы федеративной системы МО каждой стороне необходимо локально развернуть модуль обучения и передать ему данные. Модули обучения различных поставщиков данных обмениваются параметрами с использованием различных методов шифрования, обеспечивая совместное использование данных в отсутствие обмена первичными данными, что делается для защиты персональных данных. Вычислительная платформа помогает инициировать обновление модуля обучения для каждого участника и координирует взаимодействие между различными сторонами.

Техническая архитектура представлена на рисунке 4.

Техническая архитектура федеративной системы МО в децентрализованном режиме в основном формируется из вычислительной платформы, получателя результатов, инициатора задачи и нескольких поставщиков данных.

Рисунок 4 — Техническая архитектура федеративной системы МО в децентрализованном режиме

10.2 Функциональные компоненты федеративной системы МО в децентрализованном режиме

Основу вычислительной платформы составляет управляющий модуль, который распределяет вычислительные задачи между различными поставщиками данных и осуществляет координацию обучающих модулей поставщиков данных. Затем модули обучения поставщиков данных обмениваются между собой случайными числами и/или зашифрованными параметрами с целью выполнения операции федеративного МО.

Поставщик данных в основном обладает локальными данными и модулем обучения. Имеется как минимум два поставщика данных.

10.2.1 Данные

Данные поставщика данных предоставляются локальному обучающему модулю для целей МО. Поставщик данных обеспечивает отсутствие утечек имеющихся у пользователей персональных данных.

10.2.2 Модуль обучения

Данный модуль используется для получения задач МО, выдаваемых вычислительной платформой, и для выполнения операции федеративного МО на основе локальных данных, а также случайных чисел и/или зашифрованных параметров, которыми обмениваются между собой поставщики данных.

Инициатор задачи отправляет на вычислительную платформу запрос на выполнение задачи.

Получатель результата получает зашифрованный результат от поставщиков данных и, расшифровав его, получает окончательный результат.

10.3 Процедуры обработки федеративной системы МО в децентрализованном режиме

В децентрализованном режиме поставщики данных выполняют операции федеративного обучения, обмениваясь на основе установленного протокола безопасных многосторонних вычислений информацией, которая не приведет к утечке имеющихся у пользователей персональных данных.

Рекомендуемые процедуры обработки для децентрализованного решения представлены на рисунке 5.

Рисунок 5 — Процедуры обработки федеративной системы МО в децентрализованном режиме

Процедуры технической обработки в функционирующей в децентрализованном режиме федеративной системе МО включает следующие шаги:

- шаг 1: каждый поставщик данных скачивает программные инструменты с вычислительной платформы;
- шаг 2: каждый поставщик данных развертывает скачанные инструменты и подготавливает локальные данные;
- шаг 3: инициатор задачи инициирует вычислительные задачи через вычислительную платформу;
- шаг 4: вычислительная платформа координирует вычислительные задачи, передавая их на исполнение поставщикам данных;
- шаг 5: поставщики данных используют локальные данные для вычислений;
- шаг 6: поставщики данных обмениваются случайными числами и/или зашифрованными параметрами посредством множественных взаимодействий на основе протокола безопасных многосторонних вычислений;
- шаг 7: поставщики данных выполняют расчет плотных слоев (dense state) нейронной сети и получают соответствующий результат;
- шаг 8: поставщики данных передают результаты расчета плотных слоев нейронной сети получателю результата;
- шаг 9: получатель результатов объединяет результаты расчетов плотных слоев и получает окончательный результат.

Приложение А
(справочное)

Варианты использования федеративной системы машинного обучения

Федеративное МО может применяться в сервисах различных типов, которым необходимо совместное использование данных для обучения федеративных моделей и которым также требуется обеспечить безопасность данных.

В настоящем приложении описаны варианты использования федеративной системы МО с целью проиллюстрировать их концепцию и техническую архитектуру.

А.1 Пример использования: совершенствование модулей для выявления мошенничества в телекоммуникационных сетях с использованием данных из нескольких сетей

Своевременное автоматическое выявление и предотвращение мошенничества в сфере электросвязи является полезной и важной задачей. Во время общения друг с другом мошенники и их потенциальные жертвы могут находиться как в одной и той же, так и в разных сетях связи. Таким образом, для обучения модулей распознаванию мошенничества в телекоммуникационных сетях требуется тесное сотрудничество между несколькими сетевыми операторами. Однако в соответствии с действующим законодательством данные сетей связи являются защищенными и отделенными друг от друга, и раскрывать их и обмениваться ими запрещено.

Режим обучения федеративного МО может помочь преодолеть ограничения на доступ к данным нескольких сетей. В этом режиме обучения сети связи хранят у себя свои первичные данные и выгружают необходимую обезличенную информацию (например, метки данных и индекс, а не сами данные) на платформу управления моделью (см. рисунок А.1). Без ограничения общности платформа управления моделью может использовать обезличенную информацию для организации выполнения задач обучения модуля на локальных обучающих платформах соответствующих сетей связи. Задачи обучения модуля выполняются исполнителями обучения модуля на локальных обучающих платформах с использованием локальных первичных данных. По ходу обучения локальные обучающие платформы могут обмениваться и синхронизировать параметры обучаемых модулей с параметрами на платформе управления моделью.

Обученные модули для обнаружения мошенничества в телекоммуникационных сетях могут быть развернуты и выполнены в нескольких соответствующих сетях связи.

В данном случае сетевые операторы могут использовать механизмы федеративного МО и соответствующие системы для обучения и выполнения модулей для выявления мошенничества в телекоммуникационных сетях без прямого обмена первичными данными.

Общая схема выявления мошенничества в телекоммуникационных сетях с использованием данных из нескольких сетей представлена на рисунке А.1.

А.2 Вариант использования для интеллектуального управления кредитованием и рисками

Федеративное МО можно использовать в сценариях интеллектуального управления кредитованием в финансовой сфере. Зашифрованные данные кредитной истории пользователей собираются у различных сторон с целью обучения федеративной модели управления рисками, которое осуществляет доверенная третья сторона. Таким образом, поставщики данных, такие как банки, финансовые учреждения и другие взаимосвязанные стороны, могут совместно использовать федеративную модель для проверки кредитоспособности физических лиц при определении суммы кредита. Технология федеративного МО способствует интеллектуальному проведению процесса проверки и помогает снизить затраты на онлайновую проверку кредитоспособности сотрудниками-людьми. Кроме того, федеративное МО в таком сценарии может поддерживать авторизацию на стороне клиента и защиту персональных данных с целью лучшего управления рисками.

А.3 Пример использования для интеллектуального маркетинга

Федеративное МО может обеспечить точную стратегию в отношении активов клиентов и защищенную среду обмена информацией, которая способствует выявлению рисков. Например, в сценарии автострахования федеративное МО может значительно улучшить дифференцированный подход в автостраховании. Это может помочь страховым компаниям разработать усовершенствованные стратегии действий до того, как клиенты приобрели страховку, при наличии согласия клиентов. Благодаря использованию информации о владельцах автомобилей клиентов можно разделить на группы с разным уровнем риска. Кроме того, может быть проведен точный анализ репутации и рисков владельцев автомобилей, чтобы реализовать точную стратегию в отношении активов клиентов.

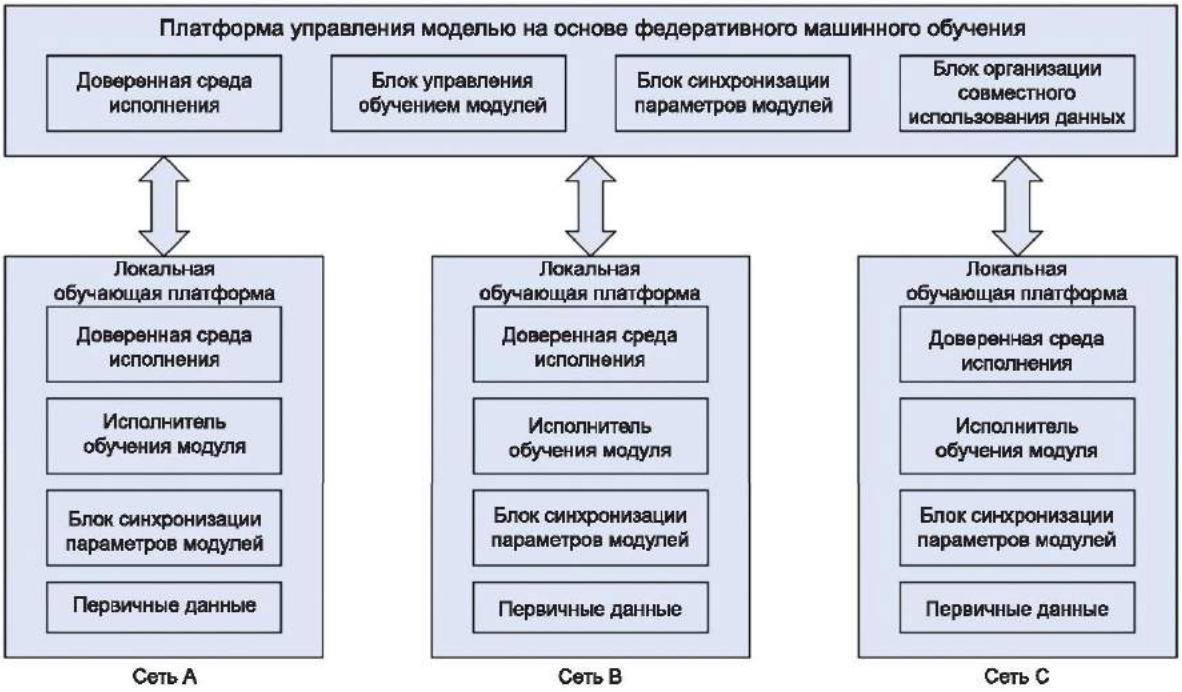

Рисунок А.1 — Общая схема выявления мошенничества в телекоммуникационных сетях с использованием данных из нескольких сетей

Библиография

[1] ИСО/МЭК 22989:2022 Информационные технологии. Искусственный интеллект. Концепции искусственного интеллекта и терминология

[2] МСЭ J.1201 (2022) Операционная система «умного» телевидения. Функциональные требования

УДК 004.01:006.354 ОКС 35.020

Ключевые слова: информационные технологии; искусственный интеллект; зашифрованные данные; безопасные многосторонние вычисления; федеративное машинное обучение; доверенная среда исполнения

Редактор *Л.С. Зимилова*
Технический редактор *И.Е. Черепкова*
Корректор *Р.А. Ментова*
Компьютерная верстка *И.А. Налейкиной*

Сдано в набор 15.12.2023. Подписано в печать 27.12.2023. Формат 60×84⅛. Гарнитура Ариал.
Усл. печ. л. 2,32. Уч.-изд. л. 2,10.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Создано в единичном исполнении в ФГБУ «Институт стандартизации»
для комплектования Федерального информационного фонда стандартов,
117418 Москва, Нахимовский пр-т, д. 31, к. 2.
www.gostinfo.ru info@gostinfo.ru