Технологии федеративного обучения для построения моделей без централизации чувствительных данных граждан

Федеративное обучение (Federated Learning, FL) — это метод машинного обучения, который позволяет обучать модели на децентрализованных данных, хранящихся на разных устройствах или в разных организациях, без необходимости обмена самими данными. Это особенно важно в сферах, где чувствительность данных имеет первостепенное значение, например в здравоохранении, финансах, телекоммуникциях, IoT и других. habr.com +2

Принцип работы

Процесс федеративного обучения включает несколько ключевых этапов:

1. **Инициализация глобальной модели.** Центральный сервер (федеративный сервер) отправляет начальную версию модели (например, нейронной сети) участникам (клиентам). habr.com +1
2. **Локальное обучение.** Каждый клиент обучает модель на своих локальных данных. habr.com +1
3. **Отправка обновлений.** Клиенты отправляют на центральный сервер не сами данные, а только обновления модели — например, изменения весов нейронной сети или градиенты. habr.com +1
4. **Агрегация обновлений.** Сервер объединяет полученные обновления для улучшения глобальной модели.
5. **Глобальное обновление модели.** Улучшенная модель отправляется обратно клиентам для следующего раунда обучения.

   Процесс повторяется до достижения желаемой точности модели.

Виды федеративного обучения

Существует несколько подходов к федеративному обучению:

- **Горизонтальное.** Модель обучается на данных из разных источников, у которых различаются записи, но совпадают признаки. Например, больницы

могут хранить информацию о пациентах разного пола и возраста с похожими симптомами болезни.

- **Вертикальное.** Модель обучается на данных из разных источников, у которых одинаковые записи, но разные признаки. Например, сотрудник может посещать разные клиники по ДМС и получать там разные услуги. Каждая клиника собирает свои данные о нём: результаты анализов или диагностику.
- **Трансферное.** Модель обучается на данных из разных источников, где частично совпадают записи и признаки. Например, пациенты могут обращаться как в государственные поликлиники, так и в частные клиники. В первых есть данные о базовых медосмотрах, а во вторых — о специализированных услугах. Модель использует совпадения в информации для обучения.

Преимущества

- **Конфиденциальность данных.** Данные остаются у владельцев, что снижает риски утечек и соответствует требованиям законодательства (например, GDPR, ФЗ 152, HIPAA). habr.com +2
- **Доступ к большим объёмам данных.** Можно использовать данные с миллионов устройств или от разных организаций, что позволяет создавать более точные и обобщённые модели.
- **Снижение затрат на передачу данных.** Не нужно пересылать огромные объёмы данных на центральный сервер, что экономит ресурсы и время.
- **Масштабируемость.** Технология позволяет эффективно обрабатывать большие объёмы данных и масштабироваться на большое количество устройств.

Инструменты для реализации

Для работы с федеративным обучением используются различные фреймворки и библиотеки, например:

- TensorFlow Federated (TFF) — фреймворк от Google, специально разработанный для федеративного обучения. habr.com +1
- PySyft — библиотека для приватного и безопасного машинного обучения, поддерживающая федеративное обучение.

- Flower — фреймворк для создания масштабируемых систем федеративного обучения.
- FedML — популярный фреймворк для федеративного обучения.
- Stalactite (от Лаборатории AI Сбера) — фреймворк для вертикального федеративного обучения.

Вызовы и риски

Несмотря на преимущества, федеративное обучение сталкивается с рядом вызовов:

- **Агрегация необъективных обновлений** может привести к искажению глобальных моделей, поэтому необходимы методы устранения дисбаланса данных.
- **Проблемы безопасности.** Существуют потенциальные векторы атак, например внедрение вредоносных обновлений или извлечение информации из обновлений модели.
- **Высокие затраты на коммуникацию** между локальными серверами и центральный сервер.
- **Необходимость специальных компетенций** от IT-специалистов для внедрения технологии.

Для повышения безопасности часто используют **дифференциальную конфиденциальность** (Differential Privacy, DP). Этот метод добавляет калиброванный шум в данные или параметры модели, чтобы скрыть индивидуальные вклады и обеспечить математические гарантии конфиденциальности. DP может применяться на различных этапах: центральная (CDP), локальная (LDP) и распределённая (DDP).

Примеры применения

- **Здравоохранение.** Обучение моделей для диагностики заболеваний (например, рака или COVID-19) на данных из разных больниц без обмена фактическими данными пациентов.

- **Финансы.** Обнаружение мошеннических транзакций: банки могут локально обучать модели на данных о транзакциях клиентов.
- **Телекоммуникации.** Улучшение качества связи, персонализация услуг, прогнозирование оттока клиентов.
- **IoT.** Анализ данных с датчиков умного дома, оптимизация работы промышленных устройств.
- **Рекомендательные системы.** Создание более точных рекомендаций на основе локальных предпочтений пользователей.

  Федеративное обучение продолжает развиваться, и его применение в будущем может стать стандартом для отраслей, где защита данных критически важна.