

АПРЕЛЬ `23

02

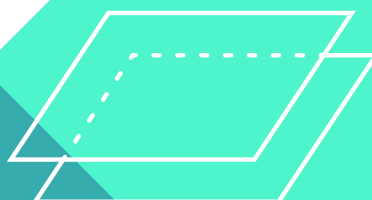


ФИНТЕХ -РАДАР

ТЕХНОЛОГИЯ МЕСЯЦА:
ФЕДЕРАТИВНОЕ
ОБУЧЕНИЕ



АФТ-инсайты
Техно-новости



БАЗА ДАННЫХ ДЛЯ БЕНЧМАРКИНГА ФИНАНСОВОГО РЫНКА

Ассоциация ФинТех инициировала создание базы данных для бенчмаркинга финансового рынка, которая делится на технологический и продуктовый бенчмаркинг.

Международные консалтинговые компании ушли из России, освободив нишу сравнительных исследований, в частности бенчмаркинга продуктовых и процессных показателей для финансовых организаций. Ранее такой работой в части продуктового бенчмаркинга занималась, например, Finalta by McKinsey, а в части технологического бенчмаркинга – исследовательская компания Gartner.

Ассоциация ФинТех запустила аналитический сервис на своей площадке, в котором регулярно будут собираться показатели и метрики по продуктам и процессам. Участники АФТ, банки, страховые компании и другие игроки российского финансового рынка получат доступ к бенчмаркам и сравнительным исследованиям.

Бенчмарки собираются по двум направлениям – продуктовое и технологическое. Направления сравнительных исследований формируются в рамках взаимодействия Управления исследований и аналитики АФТ с участниками рынка.



Рис 1. Верхнеуровневая структура создаваемой базы данных для бенчмаркинга на площадке АФТ

В предыдущем выпуске финтех-радара была представлена методология оценки уровня технологической готовности и цифровой зрелости компаний финансового рынка – исследования, относящегося к треку «Технологический бенчмаркинг¹».

¹ Данное направление включает в себя оценку уровня управления данными, информационной безопасности и противодействия киберрискам и пр.

В рамках настоящего материала производится обзор направления «Продуктовый бенч-маркинг», целью которого является сравнительная оценка уровня цифровизации, эффективности и результативности бизнеса в разрезе продуктов.

Пилотным исследованием направления стала оценка уровня цифровизации продукта «Цифровая ипотека».

В ходе работы над данным исследованием аналитики АФТ разработали методологию оценки цифровой зрелости финансового продукта по пяти направлениям:

1. цифровые каналы и клиентский путь,
2. бизнес-процессы,
3. технологии,
4. интеграция с госсервисами,
5. взаимодействие с партнерами.

По итогам пилотного исследования был разработан радар цифровой зрелости ипотечного процесса, отражающий ключевые показатели эффективности по каждому из выделенных направлений. Радар позволяет оценить текущий уровень цифровизации ипотечного процесса и перспективные направления дальнейшего развития.

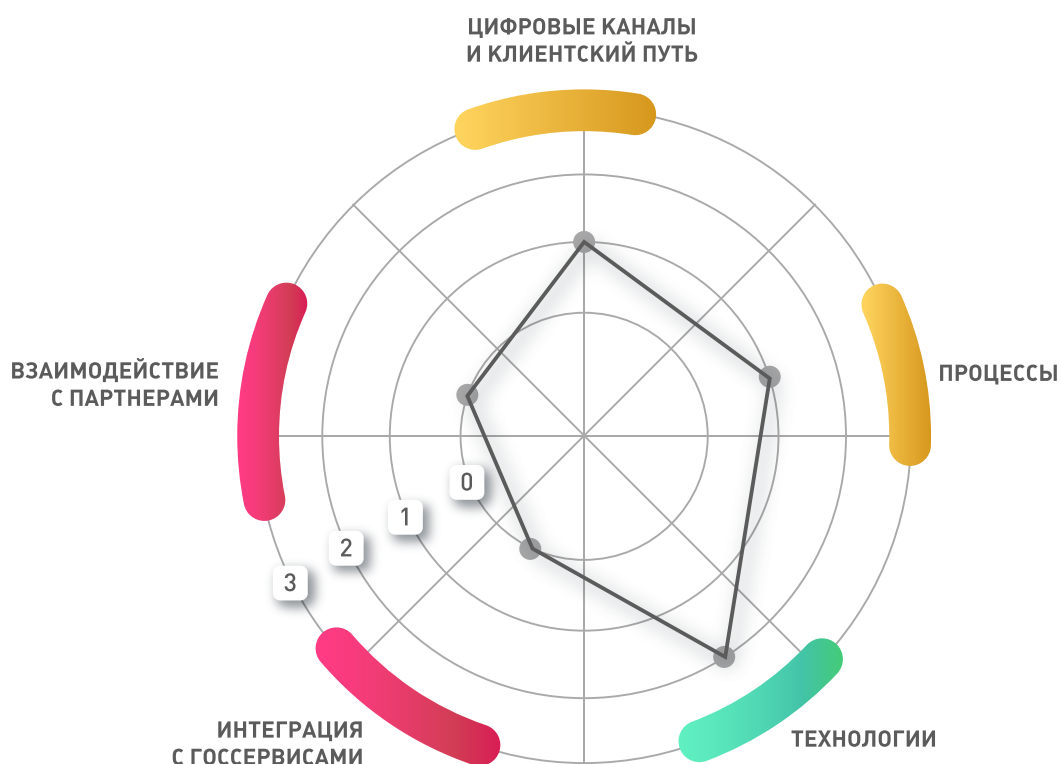


Рис 2. Пример радара цифровой зрелости ипотеки

Применение цифровых технологий в процессе оформления ипотеки позволяет банкам ускорить и автоматизировать выдачу кредитов, что улучшает качество обслуживания клиентов и повышает эффективность работы.

Отдельным фокусом в настоящий момент является применение современных технологий искусственного интеллекта для совершенствования продуктов. Использование ИИ на финансовом рынке позволяет проводить более точный анализ данных, определять тренды и риски.

ПРИМЕНЕНИЕ ИИ-ТЕХНОЛОГИЙ

Согласно результатам вышеупомянутого исследования, 83% участников внедрили роботизацию (RPA) в свои процессы в рамках ипотеки, в то время как 62% активно используют предиктивную аналитику и машинное обучение при принятии бизнес-решений.

В последние годы технологии искусственного интеллекта становятся все более популярными в различных сферах экономики, включая финансовый сектор. Рыночные игроки активно применяют ИИ для улучшения качества работы и повышения эффективности своих бизнес-процессов. **По запросу рынка Ассоциация ФинТех инициировала исследование по применению технологий искусственного интеллекта на финансовом рынке с целью содействовать внедрению ИИ в рабочие процессы банков с учетом существующих международных трендов, бизнес-кейсов и нормативной базы, которая сложилась на российском финансовом рынке.**

Аналитики АФТ разработали таксономию¹ технологий в области ИИ на основании анализа мнений участников рынка.

ЧТО ОТНОСИТСЯ К ТЕХНОЛОГИЯМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА?



Рис 3. Таксономия искусственного интеллекта

¹Принцип классификации иерархических систем

На основе данной таксономии был проведен предварительный анализ 62-х практических примеров внедрения ИИ в банках, страховых компаниях и других организациях финансового рынка. Кейсы рассматривались с точки зрения технологией и **бизнес-перспектив**.

Среди проанализированных кейсов применения ИИ в финансовых организациях почти половину составляют решения в области анализа данных (47%), затем идут речевые технологии (27%). Перспективные методы встретились всего в 2% кейсов. Это можно связать с тем, что проработанных решений по данному направлению пока нет, и что они требуют модернизации всей технологической инфраструктуры организации.

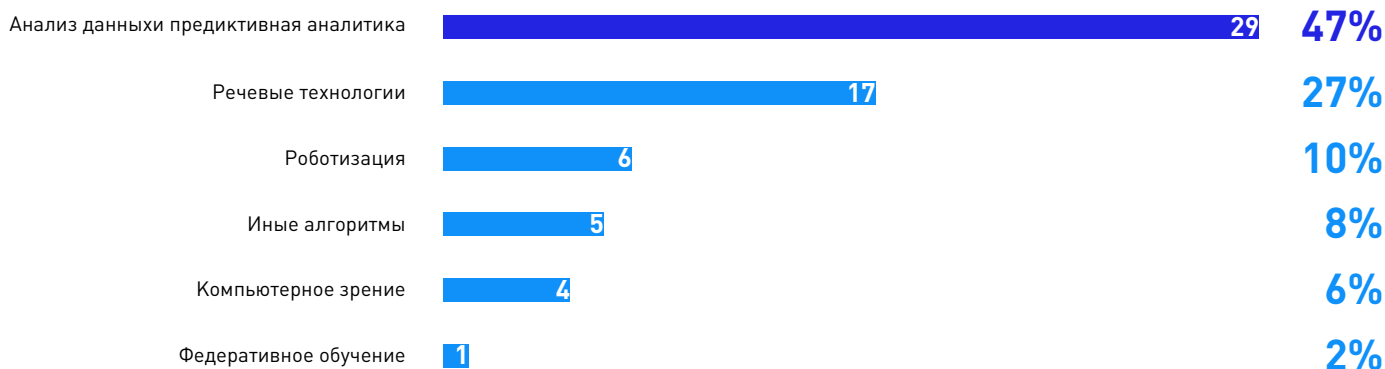


График 1. Распределение кейсов ИИ с точки зрения типов технологий

Если говорить о распределении кейсов с точки зрения бизнес-драйвера, чаще всего ИИ-решения внедряются для улучшения клиентского опыта (32%), снижения рисков (21%) и для запуска новых продуктов (16%). Это соотносится с данными по типам внедренных технологий: речевые технологии применяются для создания чат-ботов, автоматизированных колл-центров и других виртуальных агентов, а анализ данных и предиктивная аналитика в финансовых организациях используются для снижения рисков в таких решениях, как скоринг клиентов и взаимодействие с фондовыми рынками.

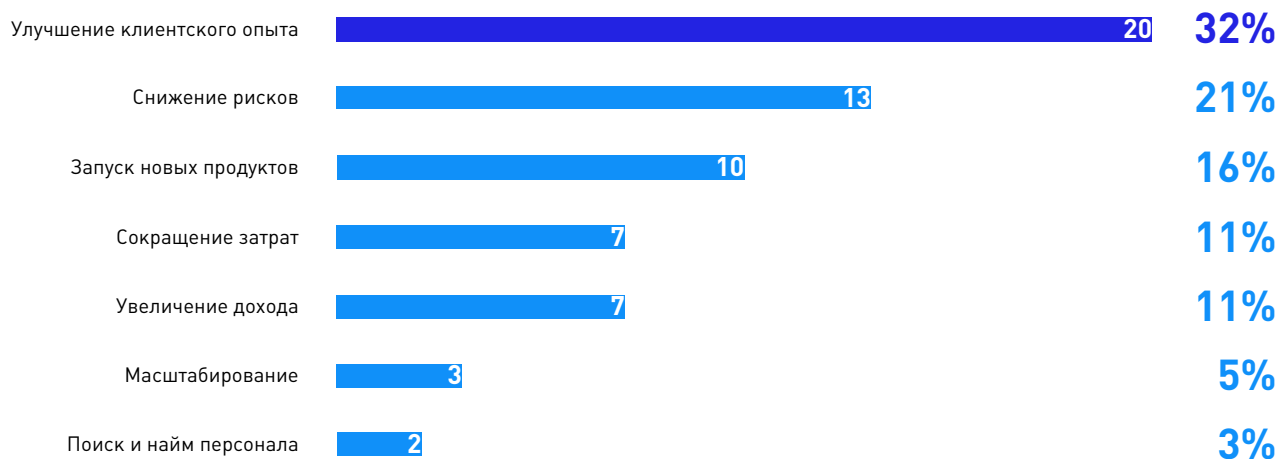


График 2. Распределение кейсов ИИ с точки зрения бизнес-драйверов

Команда исследователей АФТ провела опрос по применению различных технологий на основе искусственного интеллекта в бизнес-процессах российских организаций.



КАКИЕ ТЕХНОЛОГИИ ИИ ПРИМЕНЯЮТСЯ В ВАШЕЙ КОМПАНИИ?



График 3. Распределение применения технологий ИИ в банках и страховых компаниях

Результаты опроса показали, что российские специалисты в первую очередь используют такие технологии, как предиктивная аналитика (27%) и голосовые технологии (26%), к которым относятся распознавание и синтез речи. За ними с небольшим разрывом следуют еще два направления ИИ – компьютерное зрение, включая распознавание лиц и движений, и обработка естественного языка, подразумевающая распознавание символов и текста – их отметили 24% респондентов. Генеративный ИИ и интеллектуальная роботизация оказались внизу списка. Аналитики АФТ связывают это с тем, что оба этих направления довольно новые, поэтому еще не были повсеместно внедрены.

Аналогичное исследование проводила компания McKinsey по игрокам мирового финансового рынка. Несмотря на то, что методология исследований и рассматриваемые технологии отличаются, в результатах опросов наблюдается значительное пересечение:

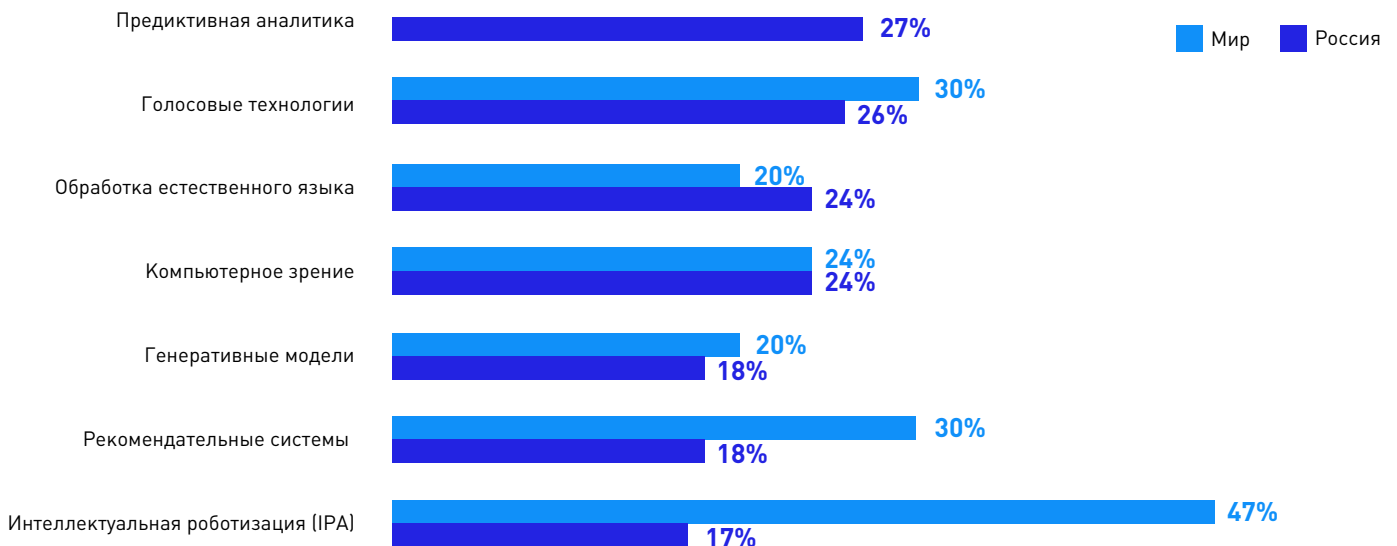


График 4. Сравнение результатов опросов АФТ и McKinsey

Анализ результатов исследований показал, что технологии обработки естественного языка более распространены в России (24%), чем по миру (20%). Компьютерное зрение используется в одинаковой степени (24%). По голосовым технологиям и генеративным моделям наблюдается небольшой разрыв в 4% в пользу мировой практики.

Наибольшее несоответствие показателей наблюдается в рекомендательных системах (18% и 30% соответственно) и роботизации (17% и 47% соответственно). Однако, в опросе McKinsey учитывалась обычная роботизация, а не интеллектуальная, в связи с чем данный показатель стоит признать нерепрезентативным с точки зрения сравнения с российским рынком.

ФЕДЕРАТИВНОЕ ОБУЧЕНИЕ



Одна из ключевых проблем, с которыми сталкиваются компании в процессе широкого внедрения технологий искусственного интеллекта, это вопрос обеспечения конфиденциальных данных. Проблема заключается в том, что при передаче данных на центральный сервер они могут быть скомпрометированы или украдены злоумышленниками. Это может привести к серьезным последствиям, таким как утечка личной информации клиентов или риск финансовых потерь. Кроме того, некоторые данные могут быть чувствительными и требуют особой защиты, например, медицинские данные или данные о финансовых операциях. Раскрытие такой информации может привести к нарушению закона или этических норм, а также к серьезным последствиям для людей, чьи данные были скомпрометированы. Для решения этой проблемы используется технология федеративного обучения.

Федеративное обучение – это подход к обучению систем ИИ, который позволяет проводить его одновременно на большом количестве децентрализованных источников данных, к примеру на смартфонах или ноутбуках, без необходимости передачи данных в централизованное хранилище.

Федеративное обучение позволяет обрабатывать конфиденциальную информацию в соответствии с законодательными нормами по защите данных и обеспечению приватности пользователей, т. к. информация в процессе обработки не передается третьим лицам. Кроме того, эта технология позволяет использовать необработанные данные, поступающие от датчиков на спутниках, мостах, машинах и растущем числе умных устройств в рамках Интернета вещей (IoT) за счет того, что их обработка может происходить непосредственно на конечном устройстве.

ПРИНЦИП РАБОТЫ

При федеративном обучении несколько сторон удаленно участвуют в обучении одной модели, последовательно ее совершенствуя, подобно совместной презентации или докладу.



Рис 4. Принцип работы федеративного обучения

Отличия в подходах к федеративному обучению заключаются в том, по каким именно данным происходит пересечение между участниками процесса. Так, между сторонами могут быть данные об одних и тех же сущностях (например, физических и юридических лицах), но разные по структуре. Рассмотрим две компании в одном городе, одна из которых банк, а другая – маркетплейс с доставкой на дом. Пересечение по пользователям обеих организаций значительное. Однако банк регистрирует доходы, расходы и кредитную историю клиентов, а маркетплейс хранит просмотренные товары и покупки пользователя – параметры их данных сильно отличаются. В этом случае применяется **вертикальное федеративное обучение (VFL)** для объединения различных параметров с сохранением конфиденциальности. Таким образом, чем более детализированной является информация, тем точнее будут прогнозы модели.

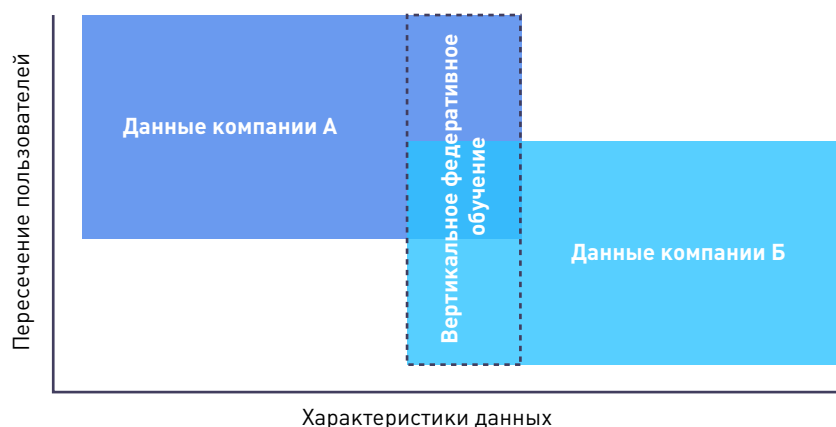


Рис 5. Вертикальное федеративное обучение

При **горизонтальном федеративном обучении (HFL)** модель тренируется на аналогичных наборах данных. Например, банки в Москве и в Тюмени предлагают аналогичные финансовые услуги, но имеют непересекающиеся группы пользователей из-за разницы в местоположении.

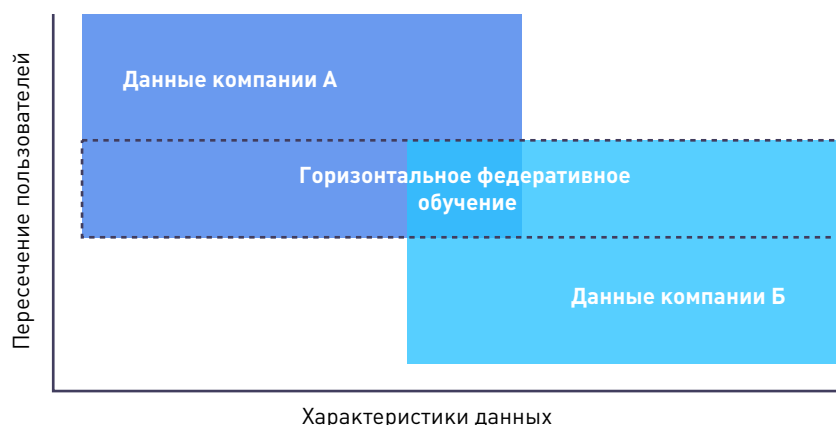


Рис 6. Горизонтальное федеративное обучение

Характеристики данных почти идентичны, а пересечение пользователей незначительно. В этом случае каждый банк обучает модели на своем объеме данных и на выходе получается модель, обученная на объеме данных двух или более организаций.

Трансферное федеративное обучение (FTL) подходит в случаях, когда у сторон почти нет пересекающихся пользователей, поскольку они работают в разных доменах и регионах.

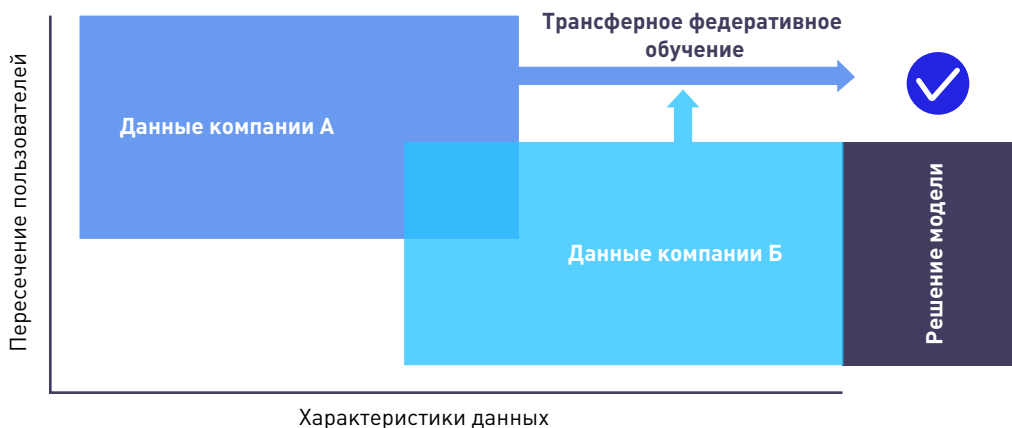


Рис 7. Трансферное федеративное обучение

Данный тип федеративного обучения заполняет недостающие метки из предварительно обученной модели, чтобы расширить масштаб доступных данных и, по сути, представляет собой комбинацию вертикального и горизонтального обучений.

ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ ФЕДЕРАТИВНОГО ОБУЧЕНИЯ

«Согласно совместному исследованию группы пакистанских и корейских ученых *Applications of Federated Learning; Taxonomy, Challenges, and Research Trends*¹, наибольшая доля исследований, в которых реализуется федеративное обучение, зафиксирована в области здравоохранения (30%). За ними следуют исследования по *Интернету вещей (IoT)* и *периферийным вычислениям (edge computing)* – 25%. На исследования по *обработке естественного языка (NLP)*, *автономным транспортным средствам*, *мобильным сервисам* и *рекомендательным системам* пришлось 10%. 5% исследований посвящены финтеху».



Денис Афанасьев,

Эксперт по искусственному интеллекту и работе с данными, Ассоциация ФинТех



Частое появление федеративного обучения в публикациях в области здравоохранения объясняется тем, что в большинстве случаев медицинские данные пациентов хранятся на серверах отдельных больниц или иных медицинских учреждений и являются наиболее чувствительными данными для конечного пользователя. Ввиду того, что передача их в единое хранилище не соответствует законодательству многих стран, то для решения задачи применения ИИ в медицине стали активно применять подходы федеративного обучения, которые выступают лучшей альтернативой сбору огромной единой базы знаний.

¹<https://www.mdpi.com/2079-9292/11/4/670>

Подход уже использовался для [сегментации мозга и опухолей](#) в медицинской томографии, а также позволил моделям работать с чувствительными данными фМРТ¹-изображений [для выявления биомаркеров заболеваний](#) в рамках Международной инициативы по обмену данными при нейровизуализации (ABIDE).

Федеративное обучение – один из основных подходов при работе с периферийными вычислениями (edge computing) и Интернетом вещей (IoT). Оно хорошо подходит для них в связи с быстрорастущей вычислительной мощностью устройств, а также необходимостью хранить и обрабатывать данные локально, соблюдая необходимые правила конфиденциальности.

Подход эффективен и при разработке рекомендательных систем, к которым относятся рекомендации товаров в интернет-магазинах, музыки или фильмов в стриминговых сервисах, а также персонализация клиентского опыта в целом. Федеративное обучение позволяет создавать системы персонализированных рекомендаций, не опасаясь утечки данных. Среди прочих FL решает проблему «холодного старта»². В случае с федеративным обучением система распознает, какой продукт или услугу предложить пользователю – профиль предпочтений нового пользователя X схож с профилем существующего пользователя Y.

ПРИМЕНЕНИЕ В БАНКАХ И ФИНАНСОВЫХ ОРГАНИЗАЦИЯХ

Федеративное обучение тесно связано с концепцией Открытого банкинга (Open Banking), Открытых финансов (Open Finance) и Открытых данных (Open Data).

Ассоциация ФинТех в 2022 году провела исследование по Открытым API в России. В нем команда аналитиков АФТ более глубоко рассматривает направления Open Banking и Open Finance, а именно изучает мировые и российские практики и проводит связь с концепцией Open Data.

Открытый бандинг — это модель бизнеса, которая предполагает открытый доступ к банковским данным и сервисам через стандартизированные интерфейсы. Основными особенностями являются:

- открытый доступ к данным о клиентах, таким как история транзакций, информация о счетах,
- появление новых сервисов, таких как аналитика расходов, персональные финансовые менеджеры и т. д.

¹Функциональная магнитно-резонансная томография

²«Холодный старт» — это ситуация, когда модели не хватает данных для корректной работы и предоставления точных рекомендаций. Она может возникнуть в трех случаях: при запуске новой системы, когда нет достаточного количества пользователей и их взаимодействий с объектами; для новых пользователей, которые еще не совершили достаточно действий для формирования их профиля; и для новых объектов, которые еще не привлекли достаточное количество интереса пользователей для формирования предсказаний.



Основные сценарии федеративного обучения в открытом банке за счет совместного обучения моделей ИИ без потери конфиденциальности данных отдельных клиентов могут быть по следующим направлениям:

1. Обнаружение мошенничества: подход может использоваться для выявления мошеннических операций путем анализа закономерностей в нескольких банках.
2. Персонализированные финансовые рекомендации: в данном случае федеративное обучение может использоваться для предоставления персонализированных финансовых рекомендаций клиентам на основе истории операций и других финансовых данных.
3. Кредитный скоринг: для разработки более точных моделей, основывающихся на данных множества банков.
4. Управление рисками: для анализа рисков и разработки более эффективных стратегий управления ими.
5. Сегментация клиентов: для сегментации клиентов на основе их финансового поведения и предпочтений, что позволяет банкам предлагать более подходящие продукты и услуги.

ЭТАПЫ СОЗДАНИЯ ФЕДЕРАТИВНОЙ СИСТЕМЫ

Ключевые шаги для создания успешной модели федеративного обучения для обмена конфиденциальными данными:



Рис.8. Этапы построения системы федеративного обучения

1. Построение доверительных взаимоотношений между заинтересованными сторонами, установка четких каналов связи и обеспечение прозрачности процессов принятия решений.

2. Определение конкретных данных для совместного использования, а также оценки масштаба проекта, целей и задач коллектива для совместного определения проблематики в рамках применения федеративного подхода.
3. Определение потенциальной пользы для каждой заинтересованной стороны, а также гарантия наличия необходимых ресурсов у каждой организации для участия. Это необходимо для соотнесения интересов и возможностей каждой стороны.
4. Определение ключевых ресурсов проекта: сотрудников, которые будут вести проект, его финансирование, бюджет для согласования состава руководства проекта, а также источников финансирования.
5. Выявление нормативных препятствий для обмена данными между учреждениями или странами, чтобы убедиться, что операции проводятся законным образом.
6. Разработка системы принятия решений, включающей в себя установление правил доступа к данным и их использования, а также определение ролей и обязанностей всех заинтересованных сторон для создания соответствующей модели управления.
7. Разработка стандартов сбора, хранения и обмена данными между учреждениями для построения структуры данных.
8. Внедрение API-протокола, который обеспечит безопасный доступ к модели или общим данным, обеспечивая при этом защиту конфиденциальности для внедрения модели федеративного обучения для обмена конфиденциальными данными.

ПРОБЛЕМЫ ПРИ СОЗДАНИИ ФЕДЕРАТИВНОЙ СИСТЕМЫ

- **Неравномерность данных:** имеются участники сети, чьи объемы данных существенно отличаются от других.
- **Неоднородность данных:** организации собирают и обрабатывают данные по-разному, что может привести к необъективности результатов модели.
- **Замедление процесса обучения:** распределение данных между разными устройствами может затруднить и замедлить процесс обучения модели.
- **Несовместимость систем:** компании располагают различным аппаратным и программным обеспечением, что может снизить производительность системы федеративного обучения.
- **Несовершенство технологического оснащения:** недостаточно продуманное построение системы федеративного обучения может привести к задержкам и низкой пропускной способности, негативно сказывающимся на точности и производительности.
- **Распределение вычислительных ресурсов:** распределение обучения на несколько устройств или организаций требует оптимизации вычислительной мощности с учетом пропускной способности сети.
- **Роли и обязанности участников:** для успешной реализации федеративной системы необходимы четкие механизмы взаимодействия и определение зон ответственности. Большой объем данных у одного участника может существенно повлиять на результаты модели. Для того, чтобы присоединение новых участников не нарушило процесс обучения, необходимо гибкое и согласованное со всеми сторонами масштабирование технических и организационных процессов.
- **Защищенность данных:** конфиденциальность – основная причина, по которой специалисты обращаются к федеративному подходу. Передача локальных обновлений модели на протяжении всего процесса обучения может привести к разглашению данных центральному серверу или третьей стороне. Хотя в настоящее время предпринимаются попытки снизить вероятность утечки за счет использования таких механизмов, как дифференциальная приватность и протокол конфиденциальных вычислений, они часто приводят к снижению эффективности системы или уменьшению точности модели.

ИНСТРУМЕНТЫ С ОТКРЫТЫМ ИСХОДНЫМ КОДОМ И FL-ФРЕЙМВОРКИ



FL-ПЛАТФОРМЫ И РЕШЕНИЯ



БУДУЩЕЕ ФЕДЕРАТИВНОГО ОБУЧЕНИЯ

Федеративное обучение позволяет обучать модель на собственных данных множеству устройств или организаций сразу. Благодаря растущей емкости памяти и вычислительных мощностей смартфонов, планшетов, ноутбуков и даже автономных автомобилей, а также растущей скорости обмена данных благодаря 5G, FL произвел революцию в машинном обучении.

Однако есть отдельные области, в которых необходимо произвести улучшения для более массового использования этого подхода. Например, стоит изучить или разработать более

совершенные методы сбора и компоновки данных. Желательно найти более эффективные методики распределения вычислительных ресурсов, так как проекты FL становятся все более масштабными. К тому же с ростом «амбиций» моделей, построенных на федеративном обучении, необходимо разработать более гибкие и понятные методики для управления затратами.

Большинство исследований с фокусом на федеративное обучение обычно охватывают область здравоохранения и IoT. Однако этот подход также может быть полезен и в других областях, таких как системы доставки еды, VR-приложения, финансы, общественная безопасность, выявление угроз, управление автомобильным движением и пр.

ПОЧИТАТЬ ДОПОЛНИТЕЛЬНО ПО ТЕМЕ



Google AI

Федеративное обучение.
Онлайн-комикс.



PriceWaterhouseCoopers

Рекомендательные системы:
Применение в сфере финансов



BCG

Борьба за периферийные
вычисления.



Яндекс

Трансферное обучение:
почему deep learning
стал доступнее



КЛЮЧЕВЫЕ ТРЕНДЫ В ОБЛАСТИ КИБЕРБЕЗОПАСНОСТИ НА 2023 ГОД

Обзор основных трендов развития рынка кибербезопасности в 2023 году от аналитиков Gartner. Аналитики определили следующие актуальные тренды: управление рисками подверженности киберугрозам; устойчивость структуры идентификации; подтверждение надежности систем кибербезопасности; консолидация платформ кибербезопасности; трансформация операционной модели безопасности; композиционная безопасность; дизайн систем безопасности, ориентированный на человека; повышение эффективности управления персоналом; усиление контроля со стороны совета директоров. 9 выделенных трендов рассматриваются в рамках трех ключевых направлений – восприимчивые экосистемы, подходы к реструктуризации методов обеспечения кибербезопасности, а также перераспределение ресурсов для обеспечения оптимального соотношения между риском и защитой информационных активов.

РАБОЧИЕ МЕСТА БУДУЩЕГО

Отчет World Economic Forum предлагает обзор эволюции рабочих мест и развития профессиональных навыков в течение последующих 5 лет. Материал предлагает анализ ожиданий работодателей, позволяющий по-новому взглянуть на то, как социально-экономические и технологические тенденции будут формировать «рабочие места будущего».

МОДЕРНИЗАЦИЯ БАНКОВСКИХ ТЕХНОЛОГИЙ

В материале Oliver Wyman банкам предлагается рассмотреть три вопроса, которые позволят определить, подходящее ли сейчас время для обновления основных систем учреждений в рамках дорожной карты по модернизации технологий – не мешает ли состояние основных систем текущим бизнес-операциям, необходима ли модернизация «ядра» для удовлетворения потребностей рынка и какие меры необходимо предпринять для подготовки учреждений к будущим изменениям.

ИНВЕСТИЦИИ В КВАНТОВЫЕ ВЫЧИСЛЕНИЯ ДОСТИГАЮТ НОВЫХ РЕКОРДОВ

Согласно прогнозам McKinsey, к 2040 году объем рынка квантовых вычислений потенциально может достичь \$106 млрд. Аналитики предполагают, что наибольший экономический эффект от квантовых вычислений получат 4 отрасли: автомобилестроение, химическая промышленность, финансовые услуги и естественные науки.





КОМАНДА ИССЛЕДОВАНИЙ И АНАЛИТИКИ АФТ

МАРИАННА ДАНИЛИНА

Руководитель Управления исследований и аналитики

[E: m.danilina@fintechru.org](mailto:m.danilina@fintechru.org)



ДАРЬЯ ПЕТРОВА

Ведущий бизнес-аналитик по исследовательской деятельности

[E: d.petrova@fintechru.org](mailto:d.petrova@fintechru.org)



ГРИГОРИЙ КАРУНАС

Бизнес-аналитик по информационным сервисам

[E: g.karunas@fintechru.org](mailto:g.karunas@fintechru.org)



TELEGRAM КАНАЛ



WWW.FINTECHRU.ORG

