

03 сентября 2021

Big Data: анализ с пользой

Машинный анализ больших данных обеспечивает принятие правильных управленческих решений, выявляет такие закономерности в поведении клиентов или в работе технологического оборудования, которые не видны аналитику, и открывает путь для разработки инновационных продуктов.

ВЫЧИСЛИТЕЛЬНЫЕ МОЩНОСТИ С ЗАПАСОМ, ПРИЧЕМ НЕ ТОЛЬКО ДЛЯ СЕБЯ

Для кредитно-финансовых организаций, которые собирают и хранят данные сотен тысяч клиентов в среднем в 40 информационных системах, извлечение пользы из огромных массивов данных — нетрадиционный, но очень перспективный источник дохода.

Более того, в последнее время все больше банков в своем стремлении стать цифровыми выходят за рамки исключительно финансовых услуг и стараются привлечь на свои орбиты множество других небанковских сервисов. При создании экосистемы цифровых сервисов, отметил Дмитрий Кулагин, руководитель центра управления данными IT-компании «Инфосистемы Джет», у кредитно-финансовых организаций появляется все больше задач, связанных с машинным обучением. Все они — распознавание естественного языка, речи и изображений; транзакционная аналитика и прогнозирование; скоринг; антифрод; риск-менеджмент; рекомендательные системы — требуют



анализа больших данных и соответственно расширения вычислительных мощностей.

В ответ на новые потребности крупные кредитно-финансовые организации — «Сбер» и «Тинькофф» — создали и ввели в коммерческую эксплуатацию суперкомпьютеры «Кристофари» и «Колмогоров». Первый стал самым мощным суперкомпьютером в России, его эффективная производительность — 6,7 петафлопс, второй был признан одним из лидеров среди суперкомпьютеров в стране.

При этом позиционировались обе машины по-разному. Тинькофф Банк рассчитывал использовать всю мощь «Колмогорова» для решения внутренних задач, связанных с обучением и переобучением нейросетей, внедрять предиктивную аналитику на основе накопленных за все время работы больших данных, а также развивать функционал своего голосового помощника. А Сбербанк, создавший «Кристофари» в интересах своей дочерней компании SberCloud, на момент ввода суперкомпьютера в коммерческую эксплуатацию объявлял о том, что он предназначен для научно-исследовательских, коммерческих и государственных организаций, работающих в различных отраслях экономики — нефтегазовой, электроэнергетике, тяжелой промышленности, медицине, телекоммуникациях, ретейле, финансовом секторе и пр. Иными словами, у крупнейшего российского банка изначально были намерения задействовать вычислительные мощности суперкомпьютера для предоставления облачных сервисов клиентам — юридическим лицам.

«В ближайшие годы ряды владельцев суперкомпьютеров пополнятся новыми именами, будет развиваться сегмент аренды ресурсов суперкомпьютеров, что сделает эти технологии доступными не только для сверхкрупного бизнеса», — считает Дмитрий Кулагин. Он добавил, что у компании «Инфосистемы Джет» есть опыт построения таких систем в партнерстве с NVIDIA.

ЗА АНАЛИТИКОЙ БОЛЬШИХ ДАННЫХ — В ОБЛАКО

Анализ больших данных актуален для любых вертикальных рынков, участники которых стремятся к цифровизации бизнеса, и кредитно-финансовые организации не являются исключением. Наоборот, банки больше, чем, скажем, ретейлеры, знают о своих клиентах, а значит, заинтересованы в том, чтобы эти знания монетизировать. Однако далеко не каждый из них может позволить себе инвестировать в создание суперкомпьютера для повышения скорости и качества аналитики больших данных.

«Еще пару лет назад в России SaaS-услуги по анализу Big Data предлагались только западными компаниями, например Microsoft или Oracle», — прокомментировал Денис Кулагин, но уже в 2020 году свою платформу начали предлагать и российские игроки — Mail.ru и «Яндекс».

Сегодня Big-Data-as-a-Service (BDaaS) — очень востребованная услуга, отметила **Светлана Вронская, эксперт департамента аналитических решений ГК «КОРУС Консалтинг»**. По ее словам, для компаний, которые не могут содержать команду специалистов в области data science, облачные

сервисы по обработке больших данных делают доступными анализ информации и поиск инсайтов. «Одной из первых подобный сервис предложила компания МТС в MTS Cloud, — напомнила Светлана Вронская. — Еще один пример — сотрудничество “ИТ Град” с российским разработчиком решений для анализа данных Polymatica».

Сегодня достаточно много провайдеров, предоставляющих вычислительные мощности для анализа больших данных. «Многие из поставщиков таких сервисов работают полностью в облаке, — рассказал Михаил Попов, основатель финтех-платформы Talkbank. — Вам даже не нужно устанавливать какое-то программное обеспечение на свои серверы. Достаточно просто настроить интеграцию, обеспечивающую передачу данных для обработки и анализа в такие сервисы».

Однако ведущими поставщиками услуги BDaaS помимо упомянутой выше компании SberCloud с огромным запасом вычислительной мощности суперкомпьютера выступают «Яндекс.Облако» и Mail.ru Cloud Solution. «В составе продуктов этих облачных платформ есть все необходимые инструменты для построения полноценных решений по анализу больших данных, начиная с ETL (одного из основных процессов в управлении хранилищами данных) и заканчивая BI-инструментами», — констатировал Дмитрий Кулагин.

Совсем недавно к тройке российских гиперскейлеров добавился новый игрок. Компания «Ростелеком» представила свой новый IT-продукт, назначение которого — хранение корпоративных данных, их обработка и управление ими (кстати, он так и называется — «Платформа управления



данными»). Среди сервисов, которые национальный оператор связи планирует предлагать компаниям любого масштаба на этой универсальной платформе, — решения для потоковой и пакетной обработки данных, формирования озера «сырых», неструктурированных данных, построения хранилищ информации, управления качеством данных, визуализации. По мнению Светланы Вронской, такой набор опций позволит «Ростелекому» конкурировать со «Сбером» на поле облачных сервисов.

Как кредитно-финансовым организациям понять, подходит ли аналитика из облака для решения стоящих перед ними задач? Михаил Попов предлагает отталкиваться от стратегии в области данных, которые нужно обрабатывать, и целей такой обработки. Что для банка важнее: стать клиентом крупнейшего сервис-провайдера или получить более выгодные условия от узкоспециализированных поставщиков облачных услуг?

Есть еще один путь: экономить за счет выстраивания такой экспертизы внутри компании. В этом случае потребуются вкладывать средства в ее развитие, настройку аналитических сервисов и подгонку их под свои задачи. «В целом, работа со специализированными BI-решениями, в том числе по SaaS-модели, выгодна для компаний», — констатировал эксперт. Однако Михаил Попов признает, что решения, которое устроило бы всех, не существует. Все зависит от особенностей и масштабов бизнеса банка. По его наблюдению, компании, которые используют BI-аналитику на ежедневной основе, обычно развивают собственную аналитику или плотно интегрируются со специализированными сервисами.



По мере цифровизации бизнеса значение аналитики больших данных будет возрастать, поэтому вопрос затрат на эти цели и, следовательно, выбора модели ее потребления станет еще острее.

Источник: Банковское обозрение

