

IRIS - общепанковская MLOps-платформа для разработки и эксплуатации моделей машинного обучения

Архитектура технологической платформы для разработки и эксплуатации моделей машинного обучения признана общепанковской. Новая платформа моделирования снизит операционный риск и позволит быстро внедрять в бизнес-процессы модели любой сложности. Технологическим партнером проекта выступила компания Neoflex.

Реализованный проект обеспечит непрерывность процесса разработки и внедрения моделей машинного обучения с последующей их интеграцией в бизнес-процессы банка. Развернутая архитектура автоматизирует управление моделями, работа с ними ведется в отдельном безопасном тестовом окружении, осуществляется автоматический перенос изменений в промышленную среду.

Банк «Открытие» успешно выстраивает процессы по развитию команд моделистов, обеспечивая их передовыми технологиями. Для разработки моделей машинного обучения реализован доступ к корпоративным хранилищам данных, настроены готовые образы рабочих мест для команд с различным стеком библиотек машинного обучения.

Автоматизированные процессы переноса моделей из контура разработки в контур применения позволят в будущем сократить time-to-market задач по внедрению машинного обучения в процессы принятия банком решений, а также снизить операционный риск ручного переноса логики моделей между системами.

Уникальность проекта

Проект уникален тем, что позволяет обеспечить непрерывность процесса разработки и внедрения моделей машинного обучения с дальнейшей их интеграцией в бизнес-процессы банка. Развернутая архитектура автоматизирует управление моделями и осуществляет автоматический перенос изменений в промышленную среду. Данный процесс позволит в будущем сократить time-to-market задач по внедрению машинного обучения в процессы принятия банком решений, а также снизить операционный риск ручного переноса логики моделей между системами.

Описание проекта

Летом этого года при технологической поддержке Neoflex банк «Открытие» запустил платформу для разработки и эксплуатации моделей машинного обучения IRIS, которая стала общепанковской.

В основе платформы лежат Open Source продукты компании Databricks и других лидеров интеграции данных, моделей и процессов. Гибкость и масштабируемость платформы обеспечиваются за счет «контейнеров» системы Kubernetes, на которой реализована промышленная платформа банка.

Архитектура платформы IRIS включает в себя развернутый контур разработки и контур применения моделей: разработка ведется в отдельном окружении, при этом прошедшие тестирование модели могут быть переданы в эксплуатацию в любой момент практически без ручных операций. Важная особенность платформы — возможность бесшовного внедрения модели в среду применения. Раньше на внедрение разработанных моделей ML уходили месяцы, приходилось, например, перепрограммировать модель в другой системе. Концепция MLOps предполагает внедрение моделей любой сложности буквально несколькими кликами мыши.

MLOps-платформа IRIS развернута в масштабах всего банка для работы одновременно нескольких команд разработчиков. Она позволяет настраивать различное программное окружение под каждую команду. Вычислительные ресурсы гибко перераспределяются между командами разработчиков в зависимости от потребности. Специалисты могут оперировать различными технологическими стеками, применять различные языки программирования, переиспользовать настроенные интеграции с системами и подключенные источники данных. Такой подход становится конкурентным преимуществом для привлечения дата-сайентистов, разработчиков, инженеров, которые на рынке в дефиците.

Платформа, а именно контур применения, должна быть введена в промышленную эксплуатацию осенью этого года. До конца года в ее среду должны быть интегрированы кредитные конвейеры всех подразделений — розничного бизнеса, малого и среднего бизнеса, а также крупного бизнеса для возможности внедрения ML-моделей в этих сегментах по новой технологии.

Сейчас IRIS находится в опытной-промышленной эксплуатации. На контуре платформы уже функционируют модель лидогенерации и система раннего предупреждения по клиентам крупного бизнеса. Именно в сегменте крупного бизнеса платформа исторически зарождалась. До конца года банк планирует внедрить модель оценки риска клиентов розничного бизнеса и транзакционную модель МСБ.



Павел Николаев, управляющий директор департамента интегрированных рисков банка «Открытие»: «С внедрением промышленной платформы моделирования вопрос подключения нового специалиста к системе должен занимать не более одного дня. Каждая команда моделистов работает в собственной среде с динамическим расширением ресурсов для обучения модели. Также у нас появилась возможность оперативного переноса моделей из среды разработки в среду применения для различных команд. Именно за счет такой гибкости платформа была выбрана как основа для общепанковской платформы»



Дмитрий Первухин, вице-президент, директор департамента разработки учетных и аналитических систем банка «Открытие»: «Платформа моделирования была реализована на базе развернутой в банке промышленной среды контейнеров Kubernetes, которая обеспечила для платформы необходимые гибкость и масштабируемость. В банке ведутся работы по миграции данных, используемых при моделировании, на промышленный контур, что должно обеспечить платформе еще большую стабильность и отказоустойчивость»



Геннадий Волков, партнер, главный архитектор Neoflex: «Для внедрения платформы моделирования команда Neoflex применила в банке методологию и практики MLOps – объединение технологий разработки моделей ML (Dev) и эксплуатации разработанных моделей ML (Ops). Автоматизация и мониторинг данных, моделей, процессов на всех этапах эксплуатации платформы моделирования – прочный фундамент для дальнейшего развития процессов машинного обучения в банке»