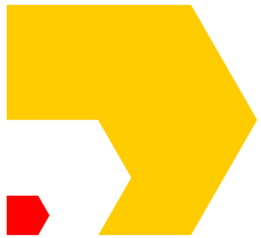


Yandex
Data Factory

Как сделать так, чтобы искусственный интеллект приносил вам деньги

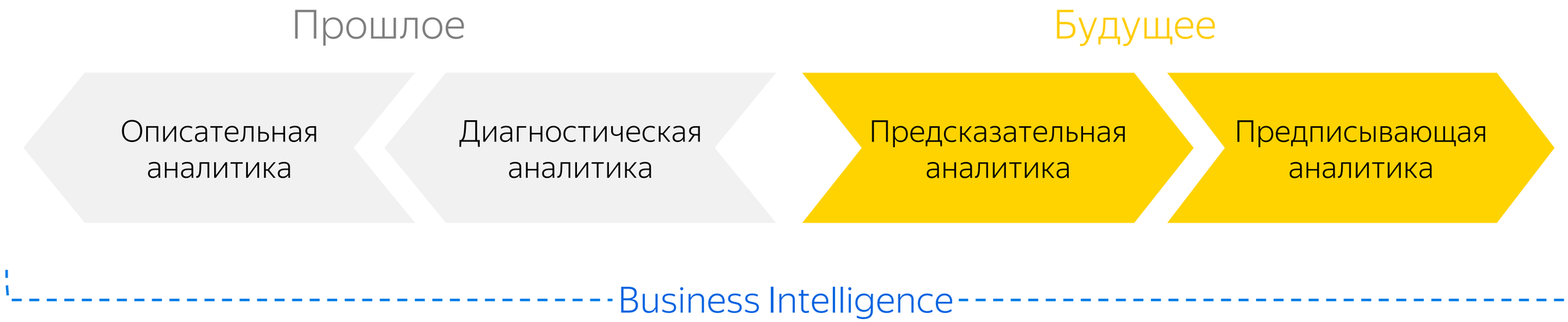
Александр Хайтин

Yandex Data Factory



- Запуск в 2014 году
- Подразделение Яндекса, ориентированное на рынок B2B
- Фокус на машинном обучении и технологиях работы с большими данными, изначально разработанными для внутреннего использования

Фокус экспертизы



Области экспертизы

- Поведенческая аналитика
- Временные ряды и аномалии
- Пространственная аналитика
- Распознавание речи, компьютерное зрение

Yandex Data Factory

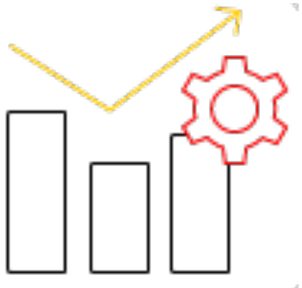
Почему сейчас?





Yandex
Data Factory

Почему Machine Learning / Artificial Intelligence?



- Во многих отраслях простые решения исчерпаны
- Дополнительная оптимизация требует капитальных затрат
- Машинное обучение и анализ больших данных позволяет извлечь дополнительную выгоду из имеющихся активов **с минимальными капитальными затратами или без них**
- С **прямым измеримым эффектом**: сокращение затрат или рост продаж

Примеры решений

Как оно бывает...

Оптимизация расхода сырья при производстве стали

Реальный кейс

МАГНИТОГОРСКИЙ МЕТАЛЛУРГИЧЕСКИЙ КОМБИНАТ
ГРУППА КОМПАНИЙ

5%

Среднее снижение расхода ферросплавов

> 275млн

ожидаемая годовая экономия

Задача

Минимизировать расход ферросплавов при производстве стали в кислородно-конвертерном цехе (ККЦ) без потери качества готовой стали

Использованные данные

- Параметры предыдущих плавок (более 200 000)
- Справочники марок сталей, ферросплавов и добавочных материалов
- Технические параметры конвертерного этапа и внепечной обработки, результаты химических анализов

Результат

Разработан программный продукт

- принимающий данные по исходному составу, массе и требованиям по содержанию хим. элементов в готовой стали
- выдающий оптимальное количество ферросплавов на каждом этапе

Yandex Data Factory

Предсказание спроса на наличные денежные средства в сети банкоматов банка из топ-10 (пилот)

Реальный кейс

Задача

Предсказать спрос на снятие наличных в банкоматах, с целью сокращения загружаемых сумм при сохранении уровня доступности. Таким образом, банк может экономить на расходах, связанными с отвлечением денежных средств.

79%

точность предсказания

Использованные данные

- Данные о банкомате (ID банкомата, адрес и тип установки)
- Данные о транзакциях (сумма, время)
- Проанализированы транзакции 4000 банкоматов за 3 года
- Данные Яндекса о погоде с высокой локализацией
- Данные Яндекса о POI (магазины, офисы)
- Данные Яндекса о событиях

Результат

- Более 40 млн. руб экономия в год на сокращении отвлеченных средств по сравнению с предыдущим решением
- 79% - точность предсказаний по всем банкоматам при сохранении уровня доступности
- Не менее 9% прироста в точности по сравнению с предыдущим решением

Yandex Data Factory

Персональные рекомендации для крупного интернет-магазина электроники

Реальный кейс

Задача

Система персональных рекомендаций и рекомендаций для блока «похожие товары» на сайте интернет-магазина

+1,5%
выручка на сеанс

+3,2%
конверсия в покупку

Использованные данные

- Логи поведения пользователей на сайте магазина (просмотренные товары, содержимое корзины)
- Ассортимент товаров по 58 регионам присутствия (название товара, категория, цена, популярность – покупки и просмотры, проводимые промоакции)

Результат

Разработаны две модели:

- Модель, определяющая вероятность покупки товара для конкретного пользователя
- Модель, определяющая степень похожести товара/ аксессуаров (для кросс-продаж)

Yandex
Data Factory

Проблемы и решения

Если все идет хорошо, значит вы чего-то не заметили...

Преимущества машинного обучения

- Новые резервы роста / оптимизации. Неочевидные решения давно изученных вопросов

- Измеримость результатов. Не нужно верить эксперту на слово, можно объективно замерить

- Минимальное time to value. За несколько месяцев можно достичь измеримого и полезного результата

- Внедрение технологий машинного обучения позволяет обеспечить существенный эффект при отсутствии капитальных вложений

Барьеры применения машинного обучения

- Отсутствие опыта использования «научного подхода»
- Необходимость культуры эксперимента (A/B тестирование)
- Риск самосбывающегося прогноза
- Невозможность персонализации ответственности
- Дискомфорт от отсутствия понимания

Что можно и чего нельзя получить от Machine Learning?

МОЖНО

- Оптимизация процессов, имеющих историю
- Рекомендации в реальном времени

НЕЛЬЗЯ

- Предсказание того, чему не было аналогов в прошлом
- Получение новых знаний, понимание
- Простые правила, убедительные объяснения

Резюме

- Использование технологий машинного обучения позволяет достичь экономических результатов без капитальных вложений. Пример ММК это подтверждает.
- Технологии машинного обучения могут встраиваться в существующие системы управления и автоматизации, позволяя получить больший эффект от оптимизации.
- Сейчас применение технологий машинного обучения является конкурентным преимуществом, скоро их неприменение может стать конкурентным недостатком.
- При правильном применении ML это не IT

Спасибо! Время для вопросов

Александр Хайтин

Исполнительный директор Yandex Data Factory



ydf-customer@yandex-team.ru



yandexdatafactory.com