



«Big Data – нефть эпохи цифровой трансформации»

Анна Племяшова
Директор по развитию
АО «Белтел»



“Data is the new oil. It’s valuable, but if unrefined it cannot really be used. It has to be changed into gas, plastic, chemicals, etc to create a valuable entity that drives profitable activity; so must data be broken down, analyzed for it to have value.”

Клайв Хамби, математик

Конференция по маркетингу Ассоциации национальных
рекламодателей 2006 г.

Что изменилось?



Развитие алгоритмов машинного обучения в последние годы, особенно методов глубокого обучения и нейронных сетей



Доступность вычислительных мощностей для быстрого обучения более сложных моделей



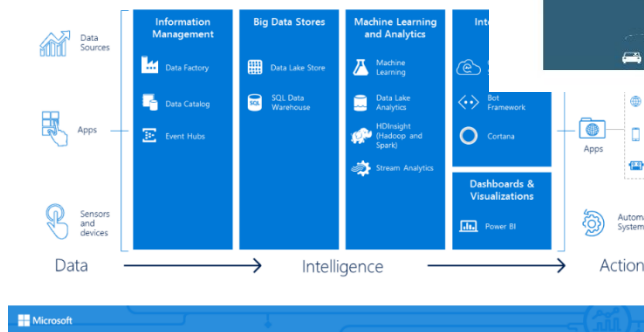
Наличие массивов данных, которые могут использоваться для обучения моделей

Что изменилось?

Преднастроенные сервисы у провайдеров облачных вычислений позволяют быстро разворачивать решения, используя алгоритмы машинного обучения

- Microsoft Azure
 - Google Cloud Platform
 - Amazon Web Services
 - SAP HANA Platform
 - IBM Watson Platform
- и пр.

Cortana Intelligence Suite Services

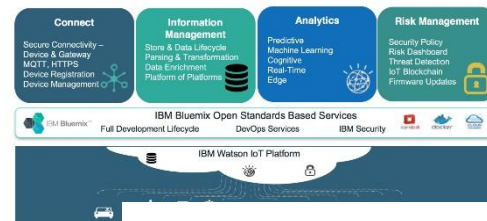


Google Cloud Platform



What is the Watson Internet of Things Platform (WIoTP) ?

The Watson IoT Platform is the environment to cognitively act upon the IoT



Что нужно, чтобы решение было реализовано или «взлетело»?



Примеры прикладных задач для бизнеса

Рекомендательный сервис



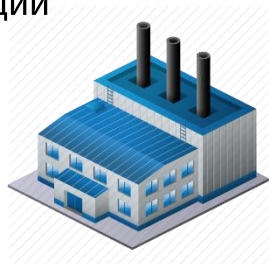
Сложности построения решения для офлайн-бизнеса:

- Усложнение идентификации пользователей, следовательно, удорожание решения
- Недостаток данных в системах лояльности бизнеса для идентификации пользователей
- Сроки реализации решения растягиваются или замораживаются в связи с отсутствием необходимых данных

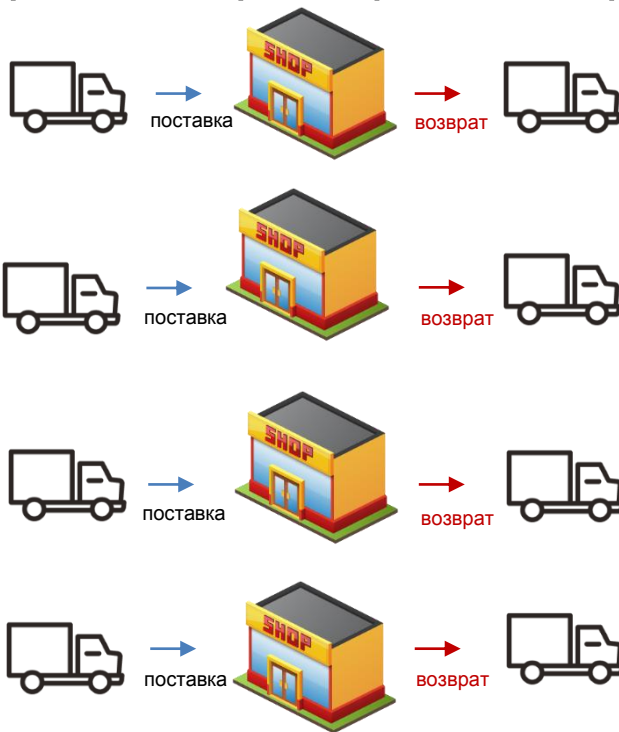
Улучшение планирования. Прогнозирование спроса и производства

Актуальная проблема – возвраты непроданного товара торговыми точками

350
наименований
хлебопекарной
продукции



Поставка в
6000 торговых
точек

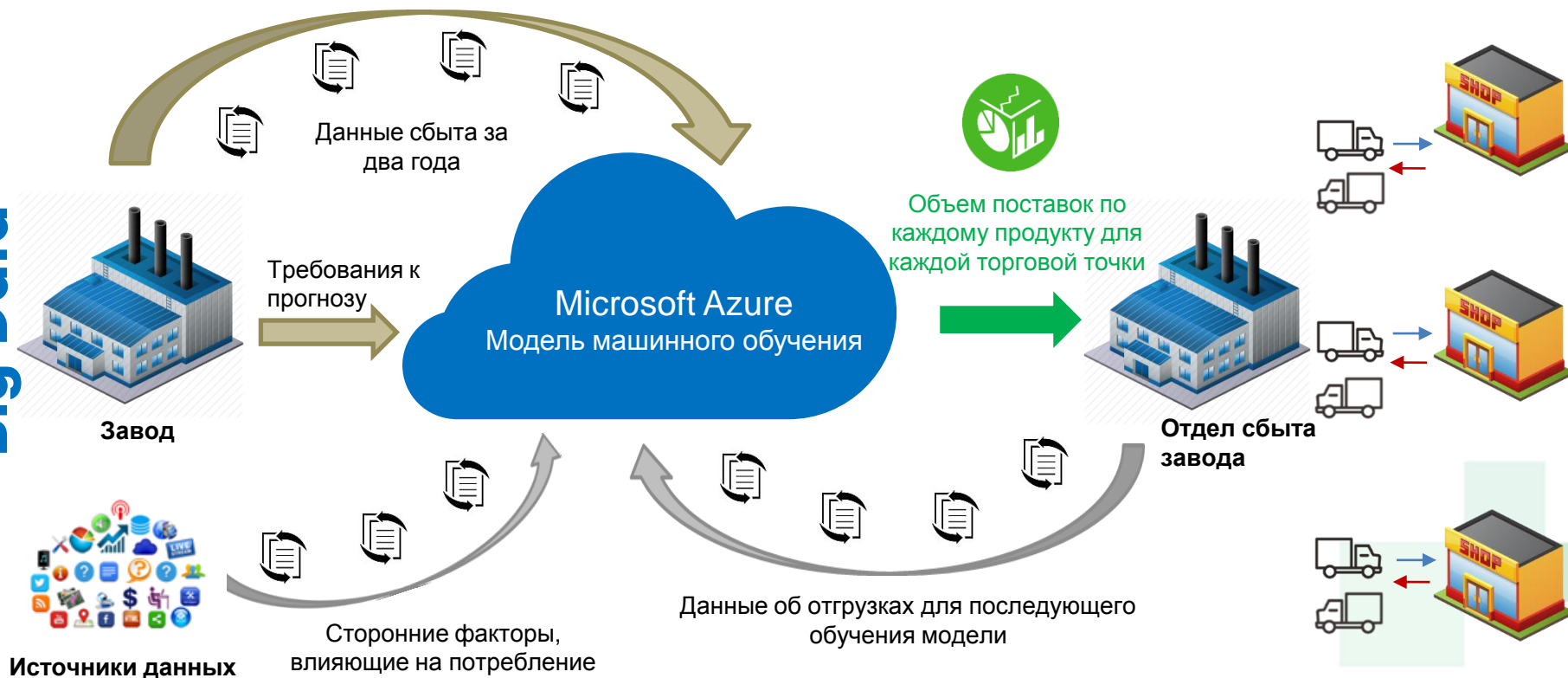


Задача: минимизировать
возвраты и, следовательно,
снизить убытки

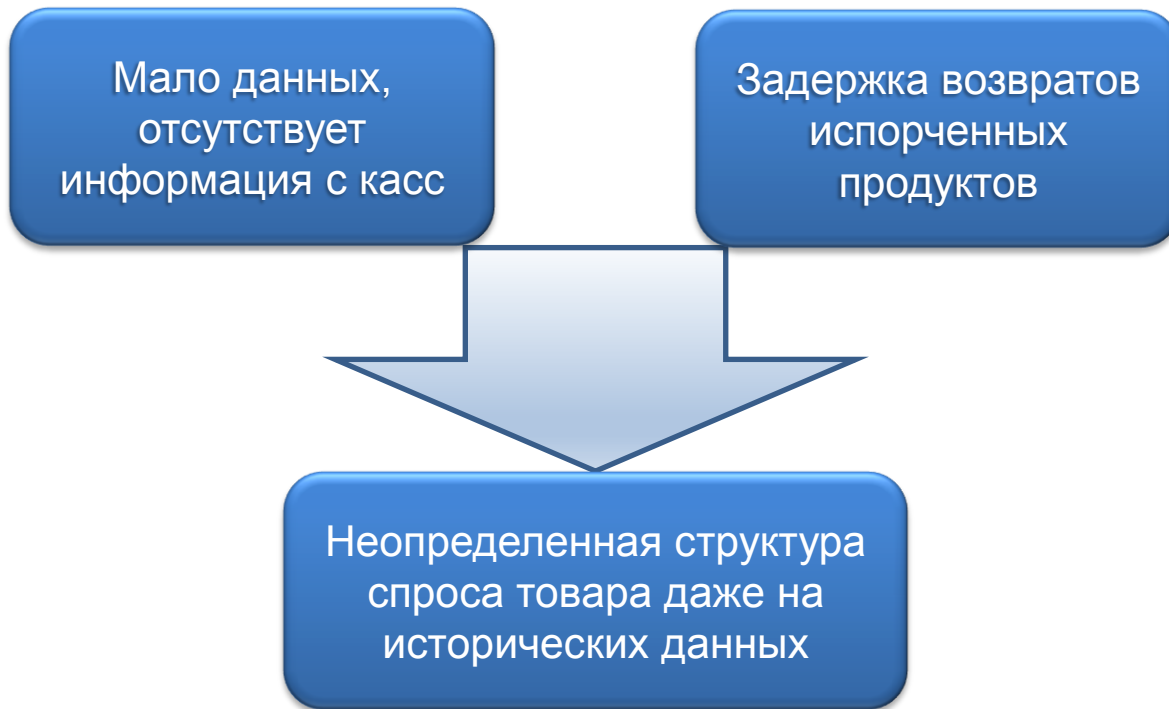
Возвраты составляют примерно 20% от объемов поставок

Прогнозирование спроса и производства. Обзор решения

Big Data



Сложности построения прогнозной модели





Точность модели

Модель приближает продажи для 80% пар магазин/товар с приемлемой точностью

График спроса хорошо прогнозируемого товара в разрезе магазина, (шт.)

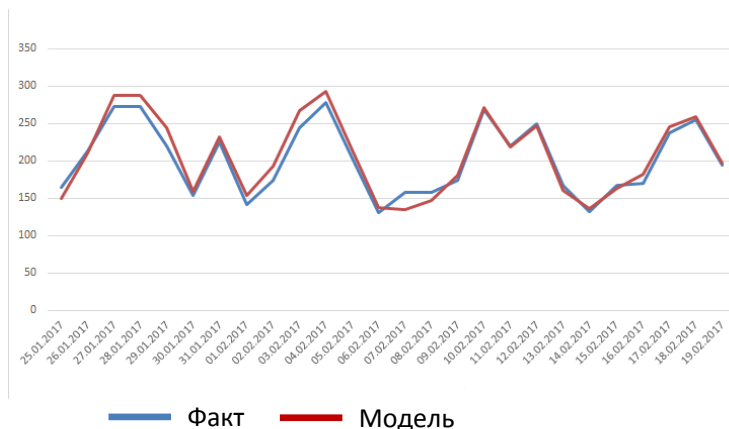
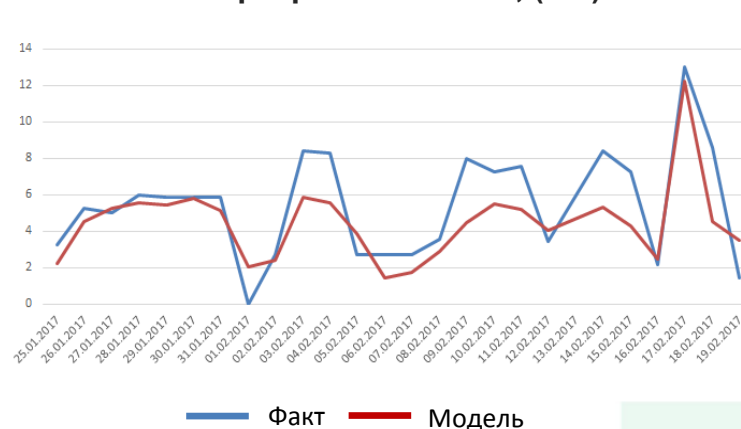


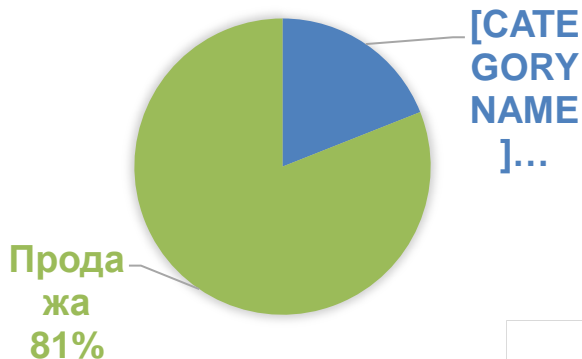
График спроса плохо прогнозируемого товара в разрезе магазина, (шт.)



Для 152 из 179 пар магазин/товар средняя абсолютная ошибка составляла не более 7%

Результат

Доля возвратов до применения предиктивной аналитики

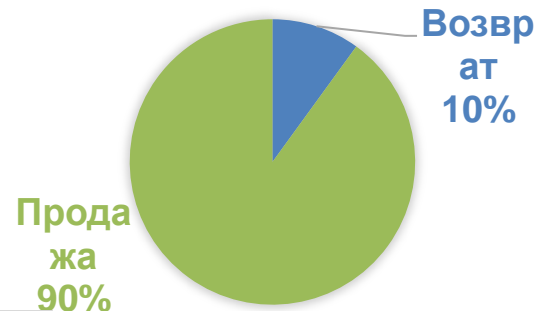


Microsoft Azure Machine Learning

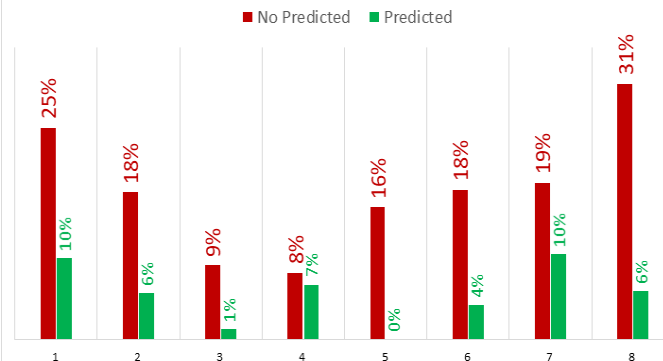


Сокращение возвратов для восьми тестовых магазинов составило 511 тысяч рублей в месяц!

Доля возвратов после применения предиктивной аналитики

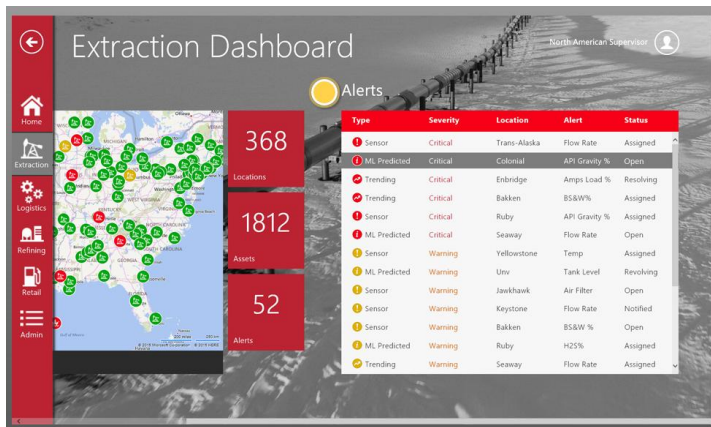


ВОЗВРАТЫ ПО 8 МАГАЗИНАМ



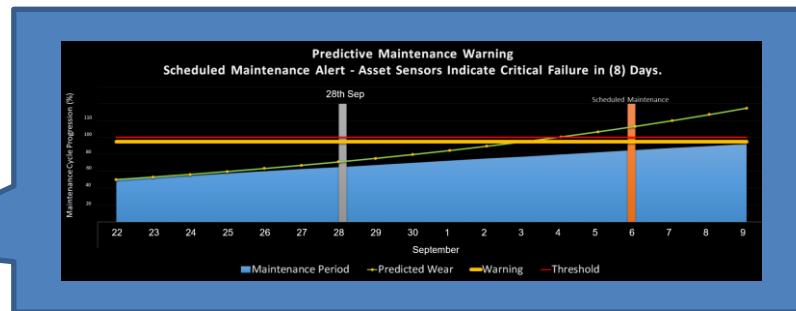
*Минимальный процент возврата необходим, чтобы исключить ситуацию пустых полок в магазинах

Повышение эффективности работы оборудования за счет предиктивного обслуживания



Результат

- **Уменьшение времени простоя** благодаря раннему предупреждению о потенциальных отказах и проблемах
- **Увеличение прибыли** благодаря возможности провести ремонт по фактическому состоянию
- **Повышение эффективности работы** благодаря возможностям удаленного мониторинга и ускорению процесса анализа неисправностей



Сложности построения решения для промышленных предприятий

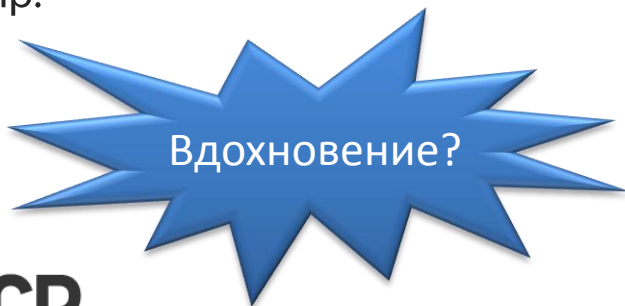
- Отсутствие необходимых данных для построения модели
- Отсутствие исторических данных достаточного срока для построение модели
- Отсроченный экономический эффект – сложность обоснования выгоды для предприятия
- Значительные инвестиции в инфраструктуру предприятия для получения необходимых данных, на которых можно строить решения



Примеры прочих прикладных задач

- Прогнозирование спроса и производства
- Формирование системы «автозаказа»
- Прогнозирование протекания технологического процесса
- Формирование цены в зависимости от характеристик объекта, динамическое ценообразование
- Прогнозирование вероятности оттока клиентов/абонентов/персонала
- Распознавание спама, мошеннических схем
- Прогнозирование вероятности невыплат по кредитам

и пр.



Парадокс цифровой экономики

Отсутствие целостных и
достаточных данных



Технологии развиваются быстрее,
и сколько бы данных не было, их
всегда будет не хватать



ДАТАНОМИКА (Datanomics) – знания, связанные с производством, потреблением, обработкой, передачей и защитой данных, получение с их помощью добавленной стоимости.

1. Юридическая неопределенность – отставание нормативной базы
 - **Ноябрь 2016** – руководитель Роскомнадзора Александр Жаров заявил, что рабочая группа по вопросам развития Интернета при администрации президента, которую возглавляет Игорь Щеголев, начала разработку законопроекта, регулирующего работу с «большими пользовательскими данными» (БПД)
 - **Март 2017** – И. Щеголев: в законе о БПД может появиться возможность запретить использование данных о себе операторам связи и другим компаниям, «у пользователя появилось больше возможностей влиять на судьбу тех данных, которые о нем получены»
 - **Июнь 2017** – А. Жаров: закон о BigData может появиться в России в конце 2018 - начале 2019 года
 - **Июнь 2017** – Фонд развития интернет-инициатив (ФРИИ) планирует разработать альтернативный законопроект о Больших данных. В его основу лягут предложения «Мегафона», МТС, «Яндекса», X5 Retail Group, ассоциации «Финтех» и других крупных компаний

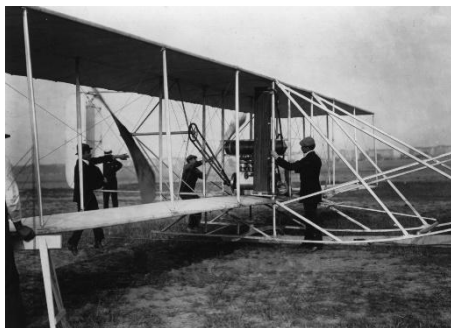
2. Ценовая неопределенность – отсутствие рынка типовых решений

- Клиенту сложно ориентироваться в адекватности предложенной цены
- Разброс цен может быть значительный в зависимости от используемого инструментария и стратегии ценообразования поставщика решения
- Часто решения по цене принимаются на уровне интуиции

3. Экономическая неопределенность – отсроченный экономический эффект решений

- На территории РФ мало доказанной практики успешных и результативных внедрений для решения бизнес-задач

Мир меняется. И с каждым днем все быстрее



106 лет



35 лет



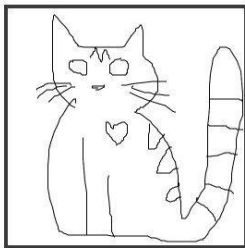
И быстрее...

edges2cats

TOOL



INPUT



undo clear random

pix2pix
process

OUTPUT



save

10 лет






Вопросы?

 +7 812 303 91 20

 datanomics@beltel.ru

 www.datanomics.ru

Подписывайтесь на канал *Datanomics* в Telegram  и страницу на Facebook

