

Data Science – стероиды для бизнеса

Никита Пустовойтов

CEO@DataUnicorn.AI



Если кратко:

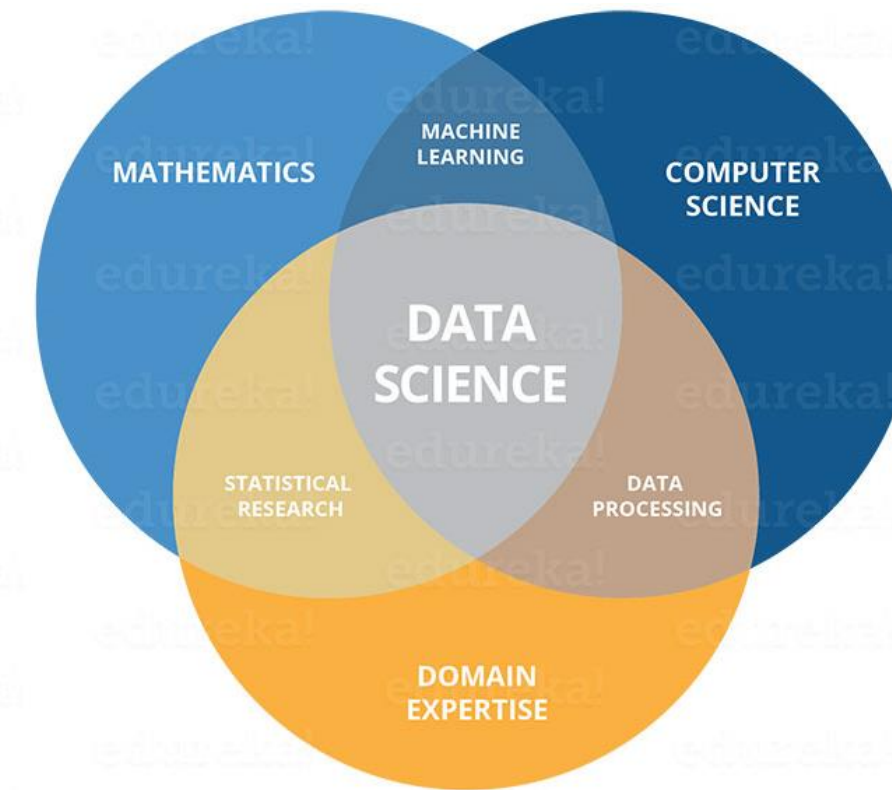
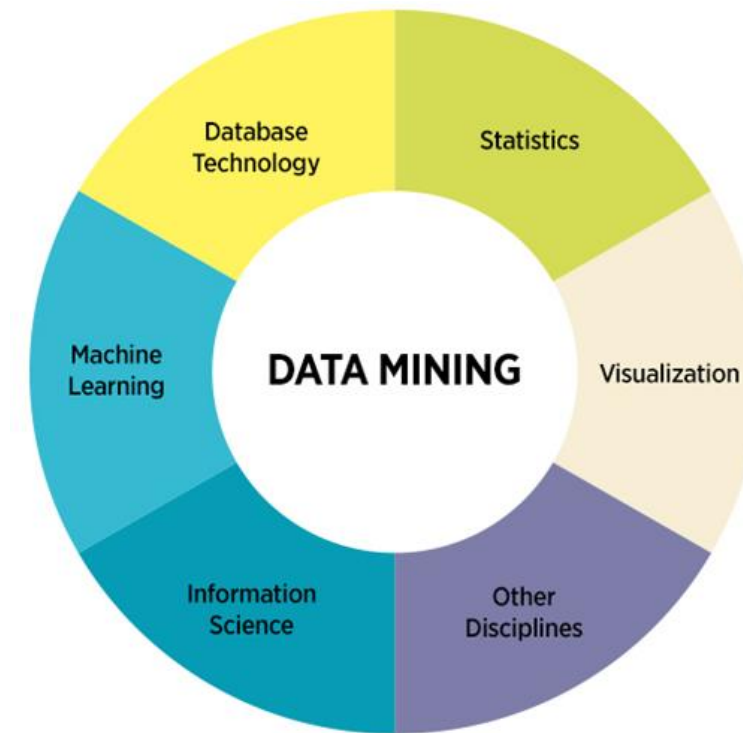
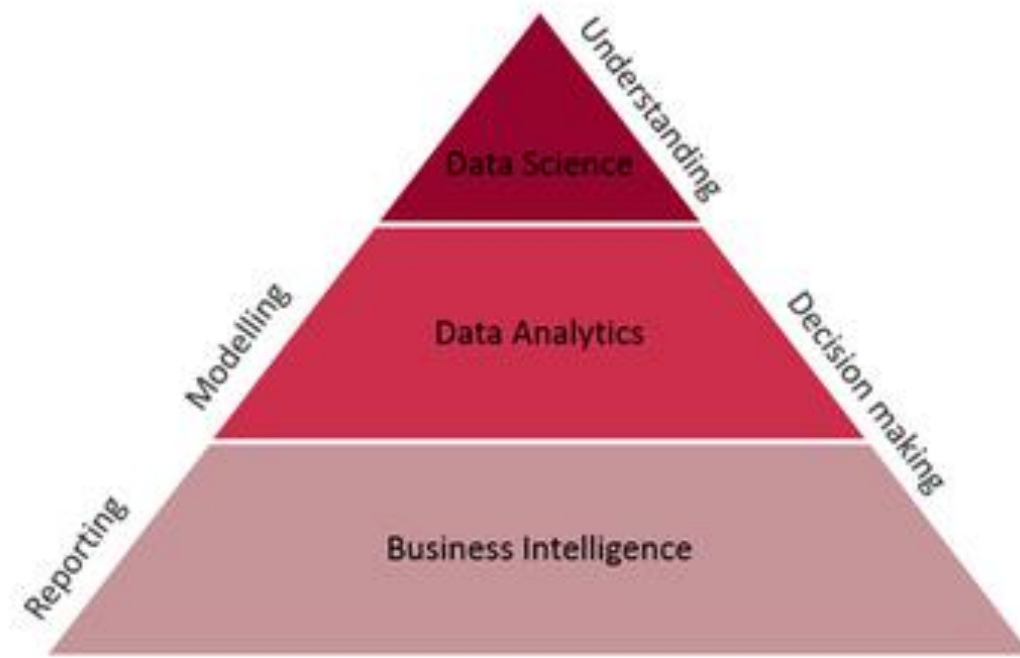
- Выпускник **ФУПМ МФТИ**, кафедра Data Mining
- **15 лет в Data Mining/ Machine Learning/ Data Science**
- 6 лет читал лекции по Computer Science на Физтехе, ставил IT-образование на ФИБТе МФТИ
- **6 лет Founder + CEO/CTO в DS-стартапах**
- За год поднимался до **278 позиции в рейтинге Kaggle**, участвуя в соревнованиях соло и тратя на соревнование 8-12 часов
- Последние **5 лет** – занимаюсь **Data Driven консалтингом** (Head of Data Science@DoubleData, Head of Data Science@BCS/CEO@DataUnicorn.AI)



Практический опыт:

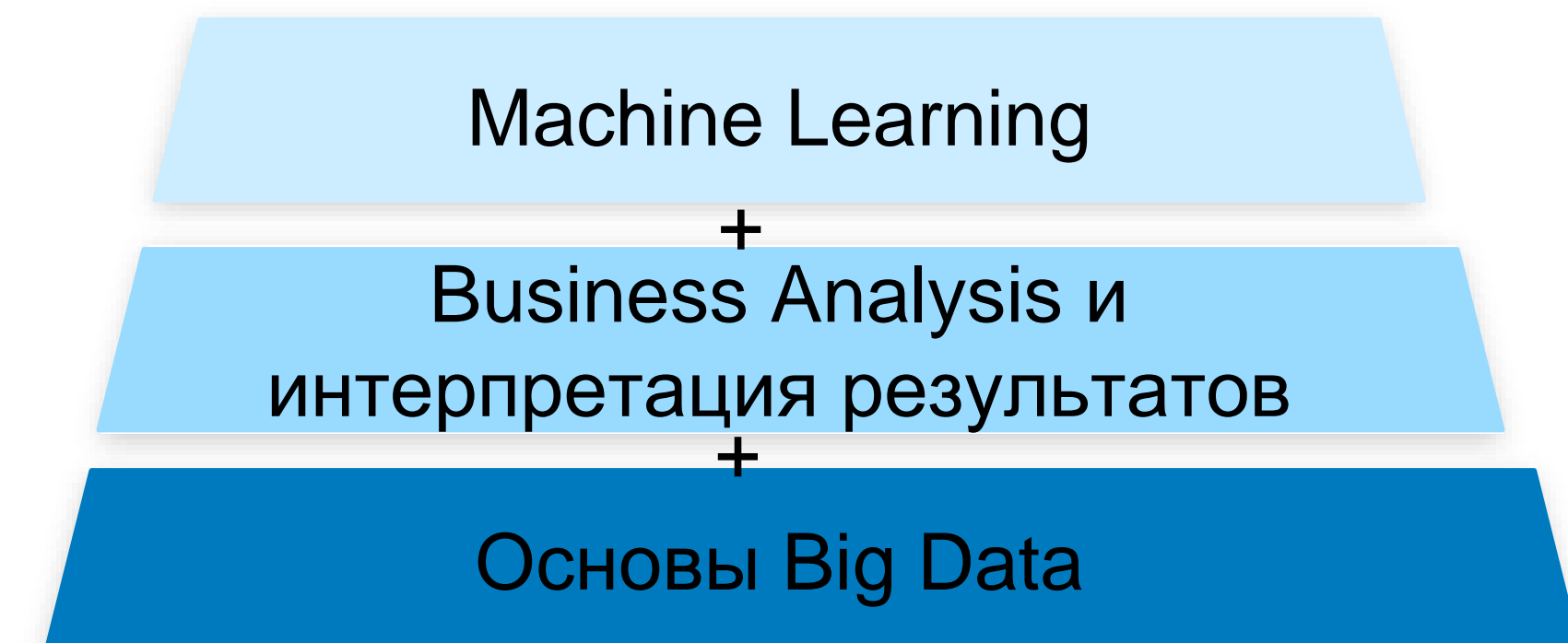
- Первый раз банковскую скоринговую систему разрабатывал **14 лет назад**
- Разрабатывал и внедрял **приложения для классификации документов/переписки**
- Прогнозировал **продажи для ритейла**
- Делал **СППР для онкологии** (диагноз, лечение)
- Применял DS в промышленности и академической науке
- Искал **фрод не только в страховании, но и в FMCG и металлургии**
- Разрабатывал **рекомендательные системы**

Что такое Data Science?

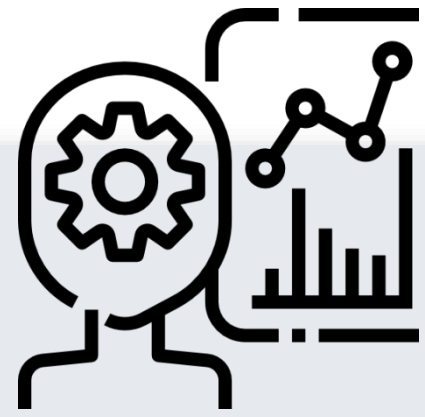


Междисциплинарная наука, использующая методы Machine Learning, мат. статистики и других областей математики и Computer Science, для извлечения знаний и закономерностей из данных с целью их применения в бизнесе

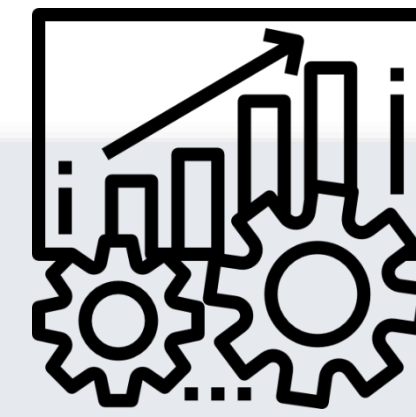
Data Science



Data Science



Фундаментальный
(*CompanyName* Research)



Прикладной

- 1 Разработка новых алгоритмов и технологий ради повышения капитализации и т.д.
- 2 Важен факт создания/обладания технологией
- 3 Приносит отложенную прибыль

- 1 Используется напрямую в операционных процессах
- 2 Приносит измеримые деньги в обозримой перспективе
 - Увеличение прибыли
 - Уменьшение расходов
 - Уменьшение Cost/Income
- 3 Часто новый источник данных и feature engineering дают больший эффект, чем авторский алгоритм

Data Science — стероиды для бизнеса

Одно из ключевых направлений для Data Science в mass-oriented бизнесе – это клиентская аналитика



DataUnicorn существует?



~ 15 человек
DS /DE/ Dev/PM



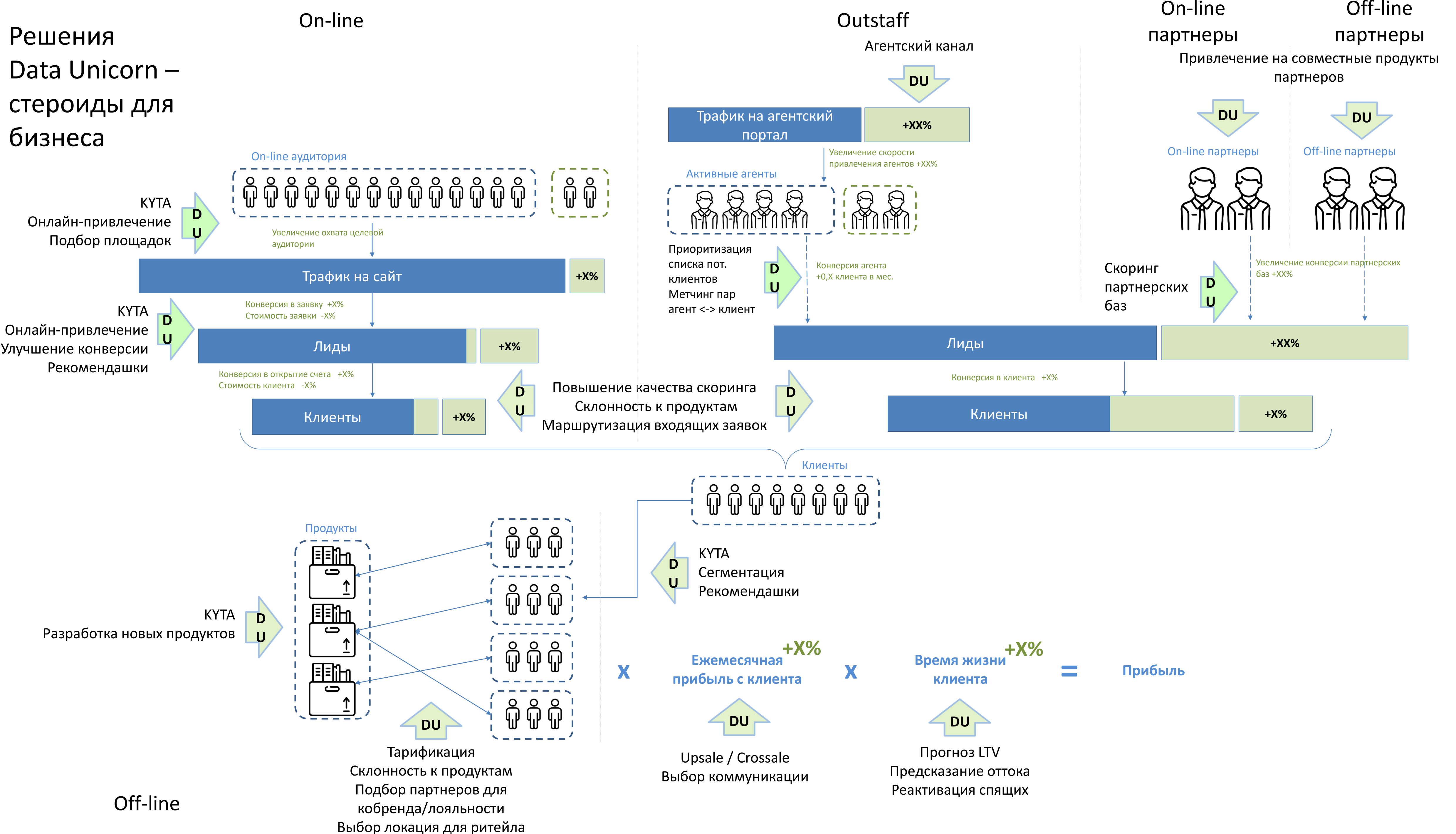
> 15 лет опыта работы с
данными в бизнесе



> 8 успешных кейсов только
за последний год.



Решения Data Unicorn – стероиды для бизнеса



Потенциальные проблемы DS в стыковке с бизнесом

«Siri вот со мной общается, а вы поток клиентов в 10 раз не можете мне увеличить с помощью искусственного интеллекта?»

(Не говоря уже о сквозной аналитике)

Неоправданные ожидания от технологии

Отсутствие базовой аналитики или технологий для тестирования

Отсутствие расчета юнит-экономики

Незрелый заказчик

Фокус только на верхней части воронки

Не всегда задачу стоит решать

Сопоставимость затрат

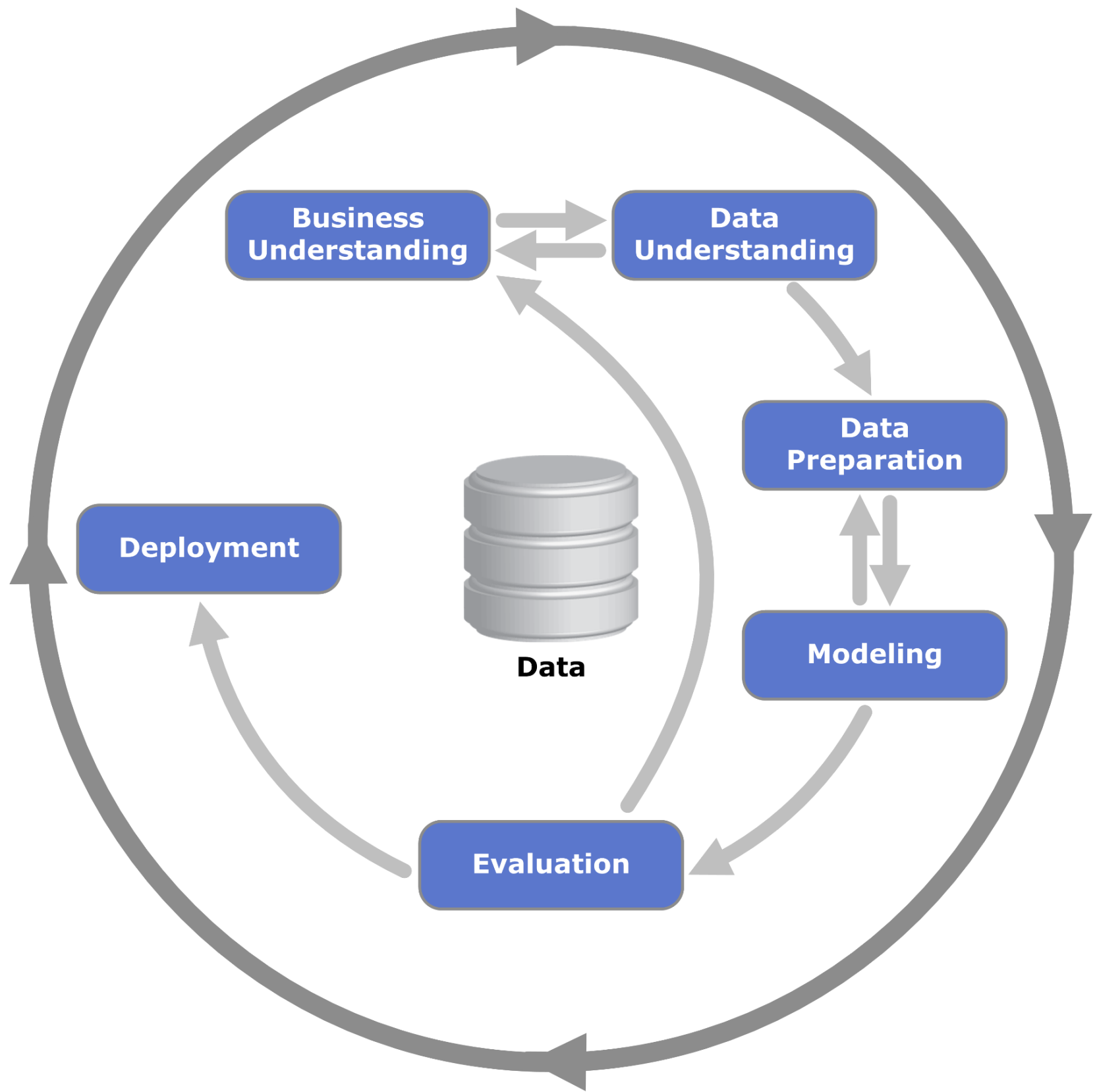
Отсутствие процессов/неготовность изменить процессы для использования результатов DS, отсутствие vision, ...

Например, нет смысла решать Churn prediction без XSell/Next Best Action


Важно оценивать TCO кейса: работа DS + затраты смежных подразделений + затраты на поддержку и сопровождение


Если объем клиентов сейчас недостаточен для сходимости экономики кейса, то кто инвестирует?

CRISP-DM – дьявол в деталях




 Обязательно должна быть связь DS с бизнесом

 Оценка: AUC-и хорошо, но надо доводить до рублей

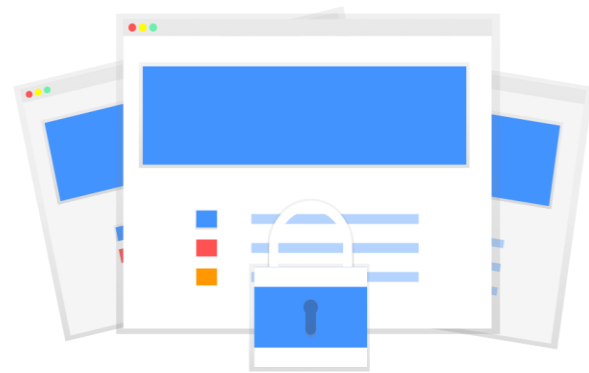
 Сравнить необходимо сравнимые вещи:
➤ Результат с моделью vs результат без модели год назад
➤ Результат с моделью vs результат без модели на клиентах

 За время проекта/ время внедрения может измениться окружающая реальность

 Контроль соответствия данных на проде тем данным, которые использовались при построении модели

Иногда Data Science может изменить модель бизнеса

1 Большой набор данных



- Кредитные рейтинги и кредитная история
- Анкета пользователя
- UTM метки
- Поведение на сайте
- Операторы сотовой связи
- Внешние провайдеры данных

2



После внедрения на Одобрим.ру **CR** из зарегистрированного пользователя в заявку **вырос на 40%**

3

По нашим оценкам, конверсия из заявки в выдачу лучше, чем у конкурентов

Данные Bankiros.ru

Конверсия	Без нас	С нами
Трафик—заявка	18,2%	23%
Заявка—выдача	2,3%	7,4%

Рост конверсий на сайте Bankiros.ru после установки виджета Одобрим.ру с Моделью принятия решений DS

4



Возможность изменения клиентского пути для части пользователей (кредит наличными -> микрозайм)

В результате изменилась стратегия, был сделан упор на White Label, стали появляться партнеры

Модель принятия решений улучшает пост-продажу кредитных продуктов



Проблема / боль партнеров

- У кредитных лидогенераторов (например, Одобрим.ру) **накапливается значительная база пользователей**
- **Воникает потребность в дополнительной монетизации старых/ холодных пользователей**
- **Присутствует непоредленность относительно продуктового предложения для данных клиентов (что им сейчас предложить?)**



Решение для партнеров

- **Текущая база пользователей разбивается на сегменты в соответствии с гипотезами**
- **Модель принятия решений определяет какой из имеющихся продуктов наилучшим образом подойдет под конкретного пользователя**



Результат

- **Конверсия из отправленной СМС в заявку ~7%**
- **Cost/ Income ratio: от 20% до 80% в зависимости от сегмента**
- **Выручка +10%** (относительно общей выручки сервиса)

Сквозная аналитика... Какой шаг следующий?

Переход на следующий этап конверсии может иметь structural bias.
Сравнивай все (и оставляй только стат. значимые результаты)!



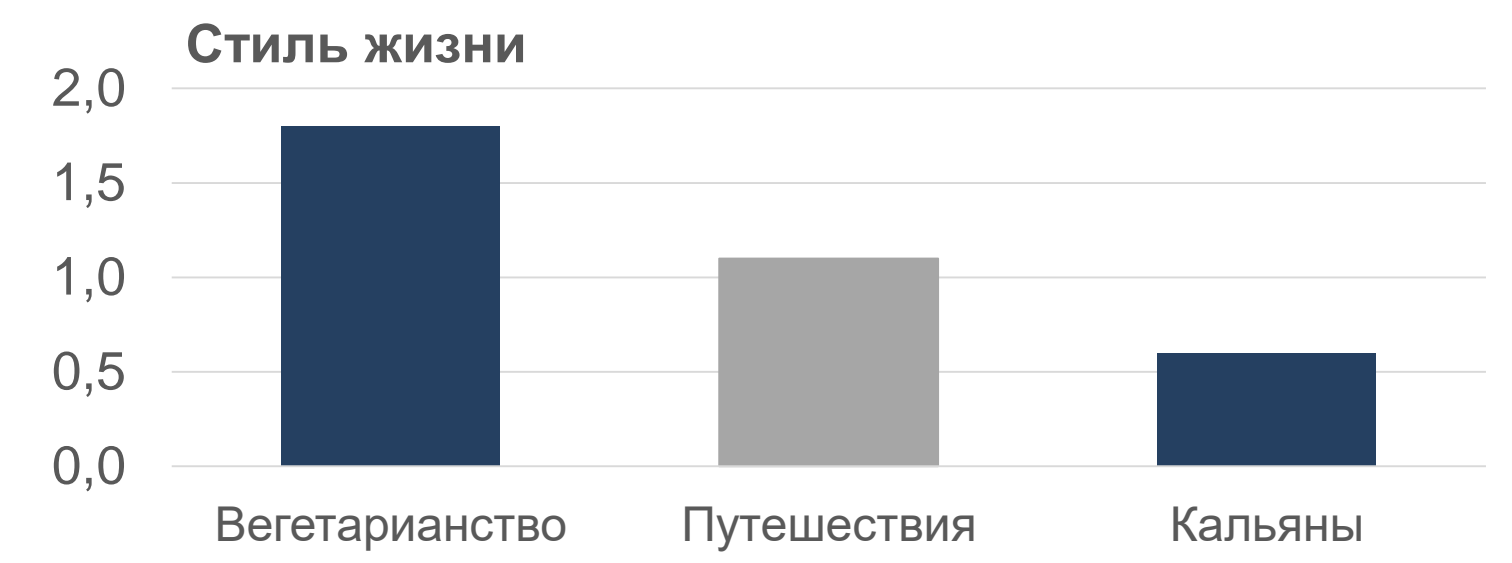
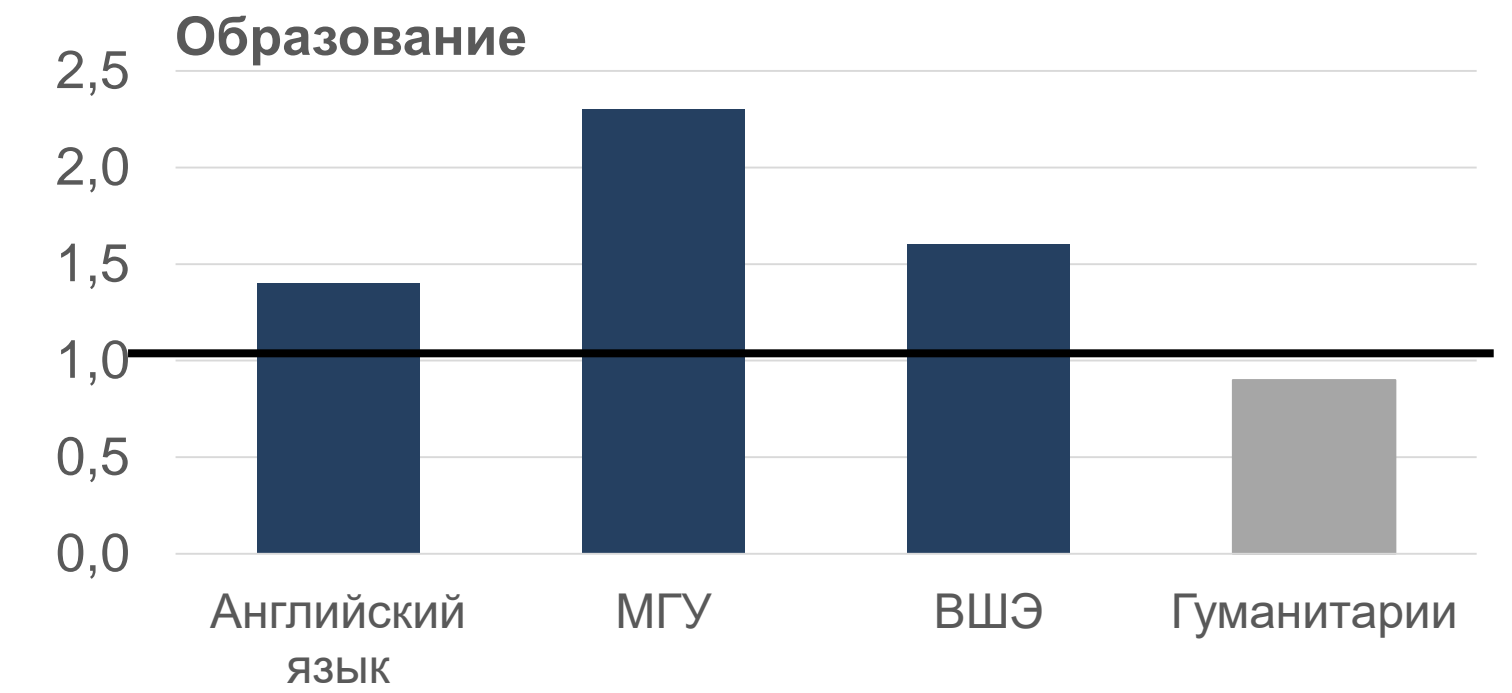
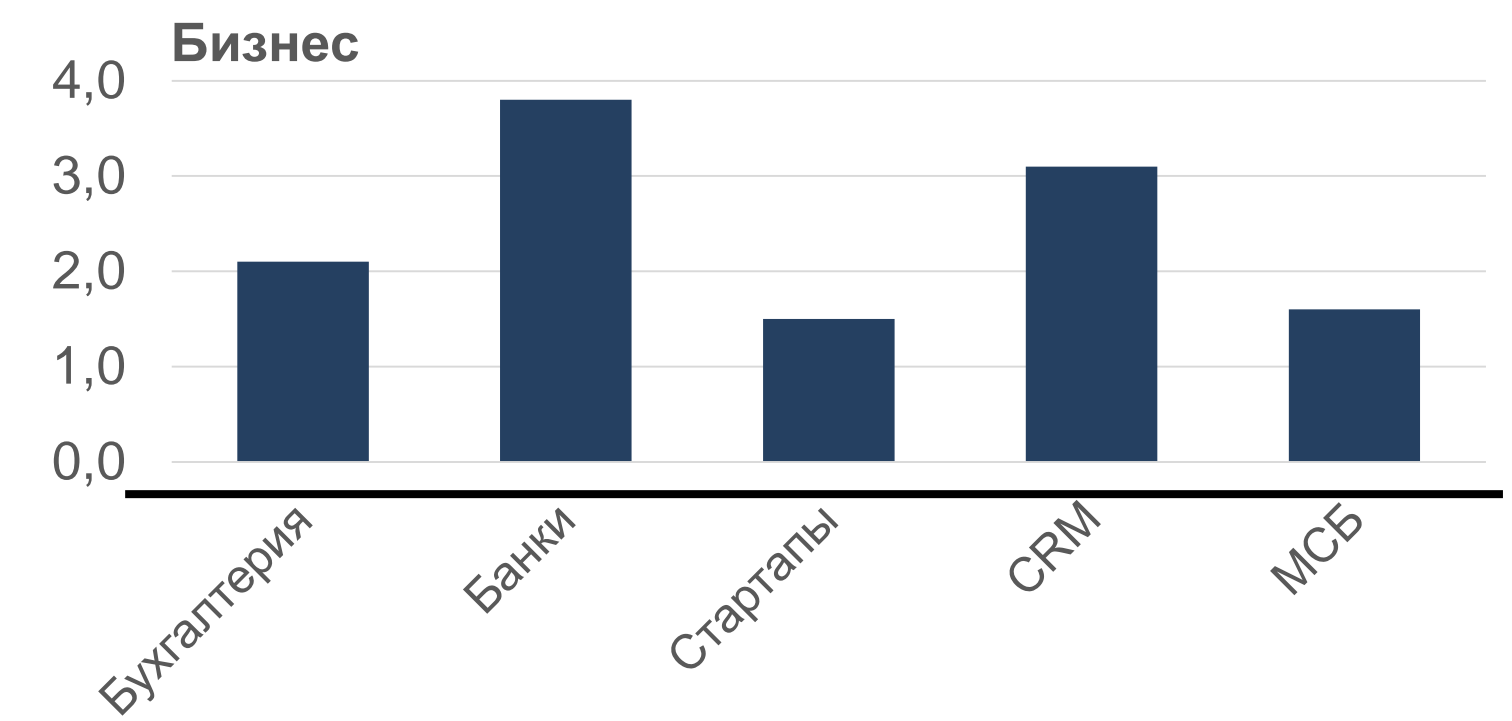
Можно использовать как внутренние, так и внешние данные



Если есть возможность – выделить внешние сегменты, сравнить внутренние сегменты с внешними, оценить емкость внешних с точки зрения потенциального проникновения и возможного дохода

Применение: маркетинговая стратегия, позиционирование продукта, сегментация аудитории, уточнение таргетингов в performance, выявление новых каналов привлечения, управление ассортиментом

Сравнение двух сегментов аудитории предпринимателей:



Откуда брать данные?

Внутренние



- CRM, анкетные данные, программы лояльности
- Транзакции
- Кликстрим на сайте
- Данные из установленного приложения
- Взаимодействие с рассылками
- Взаимодействие с ТП

Внешние



- Сотовые операторы
- Гос. данные
- Статистические данные (крупные банки, крупные отраслевые порталы, гос. статистика)
- Соцсети, приложения
- Пиксели
- Картографические сервисы
- Ритейл?
-
- DataUnicorn!

Что мы знаем о человеке?



- Пол
- Возраст
- Доход
- Семейное положение, состав семьи
- Наличие и состояние бизнеса
- Оценка наличия, марки, возраста и стоимости авто
- Оценка наличия и стоимости недвижимости
- Его/Ее интересы, хобби, увлечения
- Места притяжения

Решения для Маркетинга / CRM и Рисков



Скоринг физлиц



Анти-фрод



Предсказание дефолта



скоринг мсб



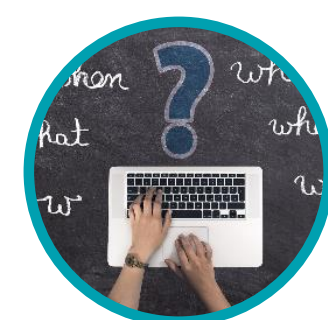
Предсказание дохода



Оффер-менеджмент



КУТА - Анализ пользователей и их предпочтений



Обогащение соц.дем переменными



Сегментация клиентов



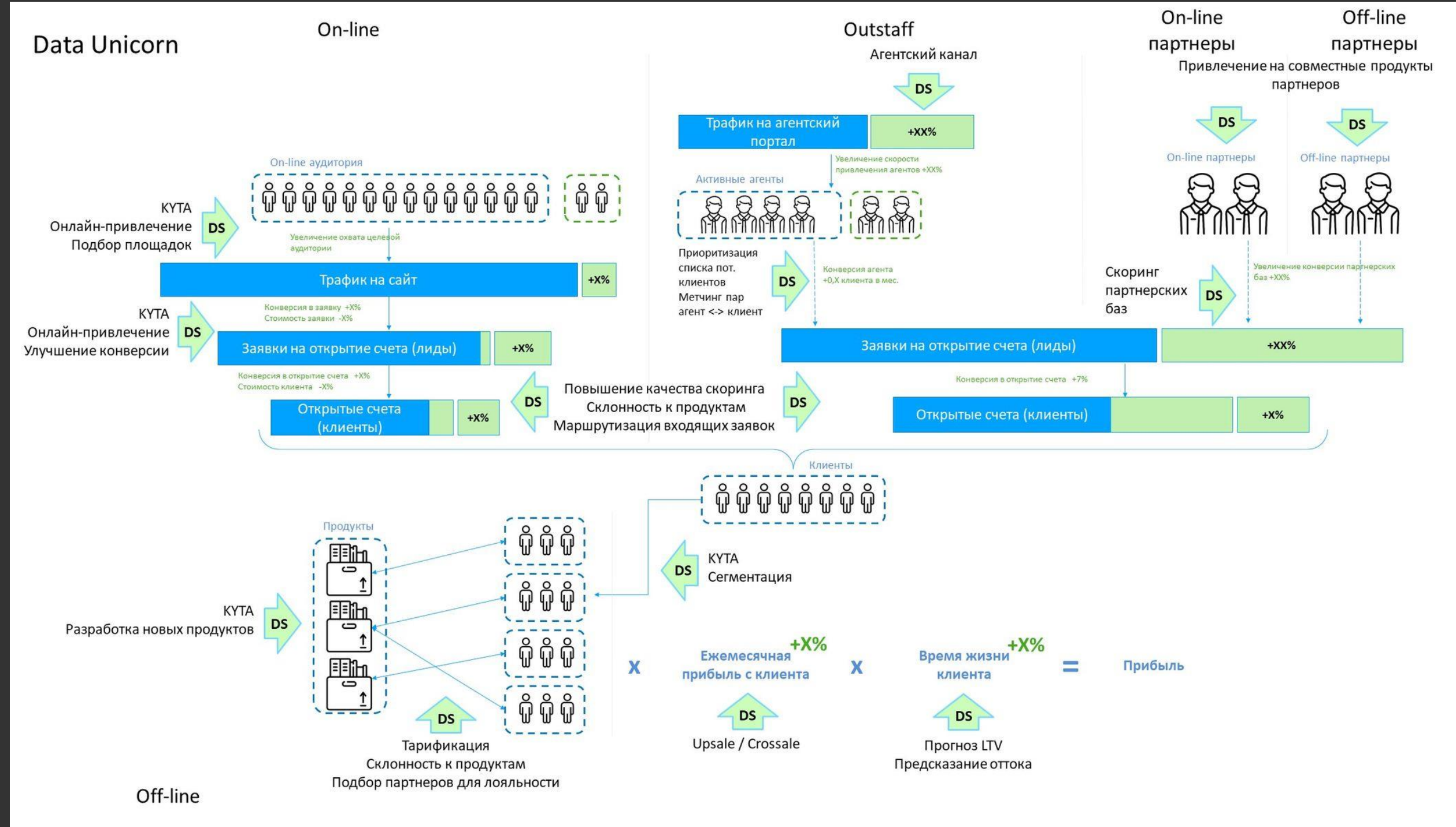
Сегменты для рекламных кабинетов



Предсказание наличия и характеристики бизнеса/авто/недвижимости



Интересы\хобби\увлечения



Спасибо за внимание!

Никита



+7 (926) 579 96 27



www.dataunicorn.ai



PustovoytovNY@msk.bcs.ru