



# БАНКОВСКИЙ САММИТ

по инновациям и развитию

КОНЦЕПЦИЯ САММИТА

**ЦФТ** ЦЕНТР  
ФИНАНСОВЫХ  
ТЕХНОЛОГИЙ

**СОЧИ '18**  
**17 · 20 мая**

Radisson Blu Paradise 5\*  
Resort & Spa

## Технологии финансовых услуг в цифровом мире

**Иван Комаров,**  
Руководитель группы  
«Машинное отделение», ГК ЦФТ

# МАШИННОЕ ОТДЕЛЕНИЕ: НАШИ ПРИНЦИПЫ

- ✓ Не хайп, а польза (увеличение доходов, уменьшение расходов, уменьшение рисков)
- ✓ Свободное ПО и сообщество дата-сайентистов
- ✓ Если можно без машинного обучения, то без
- ✓ Главное – привить культуру работы с данными и научить вопросу «Где можно выгодно применить машину?»

# ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ

- От анализа – к предсказанию – к причинно-следственным связям
- Автоматизация

## Направления

- Принятие бизнес-решений
- Снижение рисков
- Улучшение бизнес-процессов

## **ПРИНЯТИЕ БИЗНЕС-РЕШЕНИЙ: ОБОГАЩЕНИЕ ДАННЫХ О КЛИЕНТЕ**

- ✓ Вероятность оттока
- ☐ Вероятность ухода сотрудника
- ☐ Вероятность покупки
- Вероятность невозврата (скоринг)
- Рекомендации

## Вероятность оттока

- ✓ Хорошая модель – 90% точность, 90% полнота.
- ✓ Нужно смотреть дальше: предсказание строится за счет тех... кто не пользуется!
- ✓ Те кто, активный – предсказать трудно. Но некоторых можно очень точно!
- ✓ Теперь можно считать экономику – сколько можно потратить на удержание. Однако не хватает еще одного:
- ✓ Определения эффективности мер по удержанию. Делается через эксперимент.
- ✓ Получаем полностью решенную задачу: что применять и сколько зарабатываем на работе с клиентами, кто вероятнее уйдет в отток.

## СНИЖЕНИЕ РИСКОВ

- ✓ Нормализация данных
- ✓ Анти-фрод и AML
- ✓ Мониторинг операций
- Упрощение идентификации
- Оценка «нормальности» паспорта

# Нормализация данных

## Существующие клиенты:

- ✓ Оценка качества данных (хорошие, плохие, подлежащие исправлению, плохие под вопросом)
- ✓ Дедупликация (15% меняют ФИО, паспорт, имеют опечатки). Модели: поиск опечаток в ФИО, поиск по фонетическому сходству, быстрый поиск.

## Новые клиенты:

- ✓ Проверка ввода ФИО и др. данных о клиенте (абракадабра, «не похоже», правила). Модель: 99% полнота, 13% точность, срабатываний 0.17%
- ✓ Сажест (подсказки при вводе)

Оказалось важно: обучение на своих данных

## **УЛУЧШЕНИЕ БИЗНЕС-ПРОЦЕССОВ: АВТОМАТИЗАЦИЯ**

- Определение типа и страницы документа
- Определение заполнения полей документа и наличия подписи
- Разбор логов колл-центра (оценка, выделение тематик)
- Идентификация (упрощение, оценка «нормальности» паспорта)
- Голосовой робот для автоматизации рутинных операций (поиск отделений, запись на интервью)



## **МАШИННОЕ ОБУЧЕНИЕ: ВЫВОДЫ**

Сейчас:

- ✓ Подумайте и сделайте предварительную модель и начните собирать для нее данные (сделайте разметку)
- ✓ Начните проект по data science, нужно собирать корабль в плавание

Дальше:

- ✓ Возьмите существующие банковские кейсы и проведите диагностику, какие Ваши проблемы можно решить с помощью данных (это значит, что история имеет значение!)
- ✓ Если у Вас был неудачный опыт, это нормально и это всего «одно наблюдение». Можно посоветоваться с сообществом, которое такое уже делало! (Новизна, гениальность и секретность переоценены. Как и везде здесь много рутинной кропотливой работы.)

**СПАСИБО ЗА ВНИМАНИЕ!**

+7 383 336 49 49

[i.komarov@cft.ru](mailto:i.komarov@cft.ru)