Logistic Regression

> Decision Tree

Random Forest

Gradient Boosting



Социальная инженерия

Как устранить 75% fraud-потерь, не блокируя хорошие транзакции

Афанасьев Сергей КБ «Ренессанс Кредит»

Социальная инженерия



Нигерийские письма



СМС-сообщения



Телефонные звонки

Интервью со Skype-мошенником https://habrahabr.ru/post/255225/

Как зэков принуждают заниматься телефонным мошенничеством

https://meduza.io/feature/2017/06/07/polovinuotdavali-administratsii-polovinu-polozhentsu

Проблема

Растет уровень социальной инженерии с использованием переводов на web-кошельки



Решение эксперта

Выставить лимит на р2р-переводы: максимум 25 тыс. руб. в сутки

Спасем

Потеряем

Заблокируем

50%

средств клиентов 50%

средств клиентов 35%

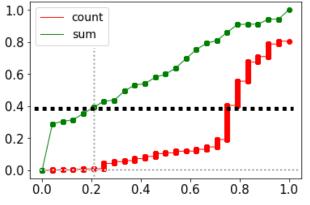
хороших транзакций (в деньгах)

А вы уверены, что это ваш клиент?

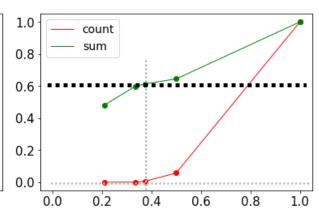


Сравнение моделей

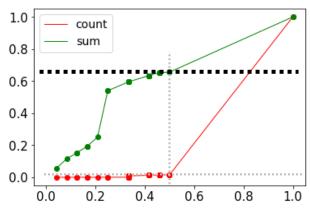




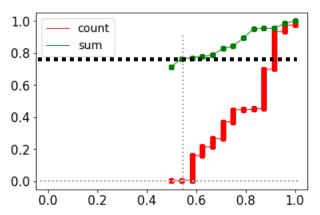
Decision Tree



Random Forest



Gradient Boosting



Выявленный fraud: 39%

Hold-транзакций: 0,94%

Затраты на звонки: 1,78%

Выявленный fraud: 61%

Hold-транзакций: 0,72%

Затраты на звонки: 1,36%

Выявленный fraud: 66%

Hold-транзакций: 1,48%

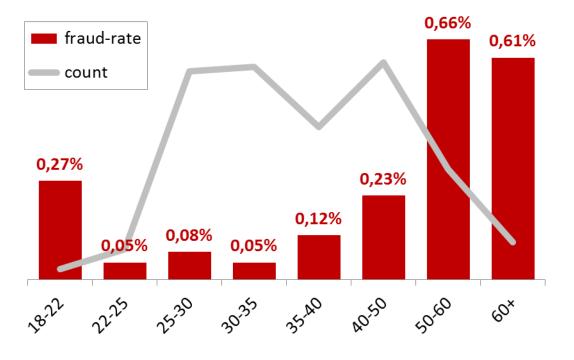
Затраты на звонки: 2,81%

Выявленный fraud: 76%

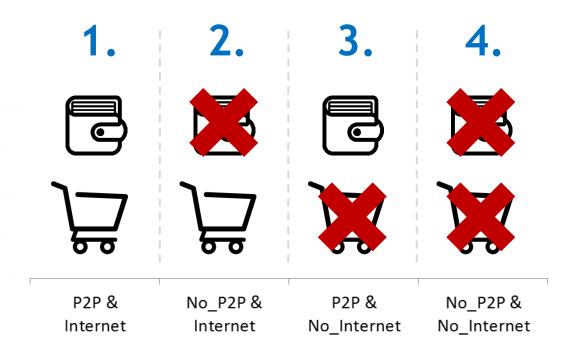
Hold-транзакций: **0,70%**

Затраты на звонки: 1,33%

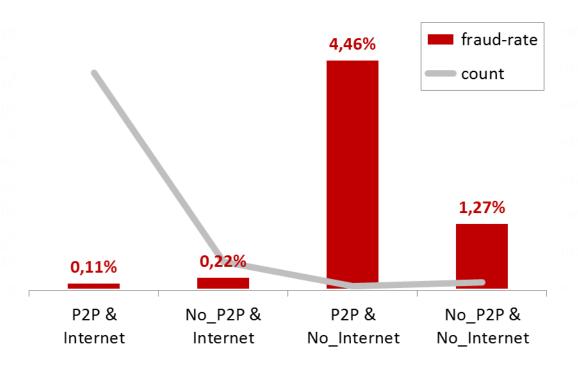




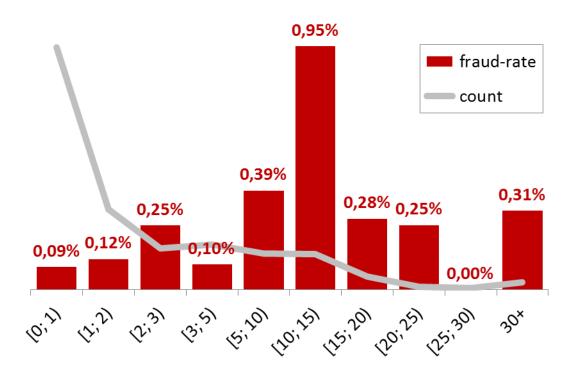












CRISP-DM:

- 1. Business understanding
- 2. Data understanding
- 3. Data preparation
- 4. Modeling
- 5. Evaluation
- 6. Deployment

Рекомендации:

Для понимания задачи и подготовки переменных необходима экспертиза антифрод-экспертов

Для антифрод-задач лучше использовать модели, восстанавливающие нелинейные зависимости: Random Forest, Boosting, Neural Network

Качество fraud-моделей необходимо оценивать с помощью Recall или F-measure

Чем меньше переменных, тем проще внедрение

Спасибо за внимание!

Афанасьев Сергей

КБ «Ренессанс Кредит»

svafanasev@gmail.com
safanasev@rencredit.ru