



команда | KB

Концепция CRM по удержанию клиентов склонных к
расторжению контракта

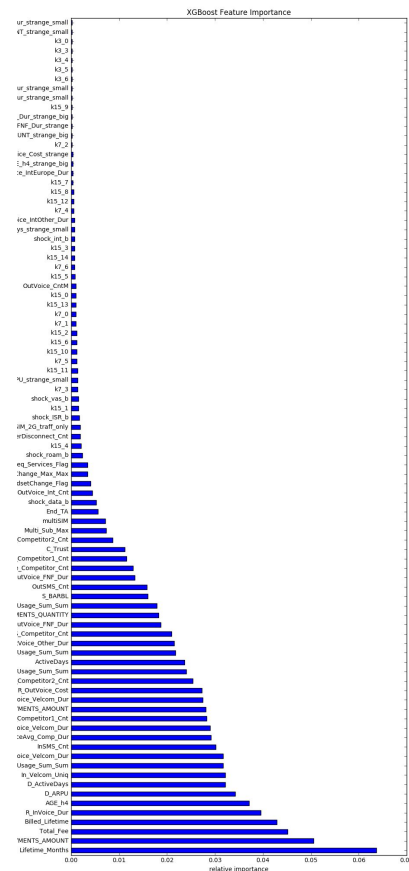


Подготовка данных

- Nan заменен медианой
- Выбросы заменены на медиану и проставлена метка о выбросе
- Нормализация от 0 до 1 (для кластеризации)

Отбор предикторов

Отброс фич с малым весом не повлиял на качество. Использованы все фичи + добавлены факторы построенные на кластеризации.

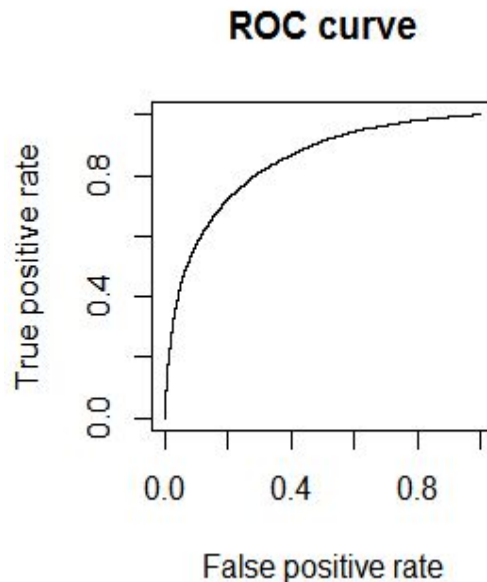


Почему алгоритм XGBoost

- Топ алгоритм для решения для подобных задач на www.kaggle.com за последние пол года
- Интерпретируемость модели принятия решений

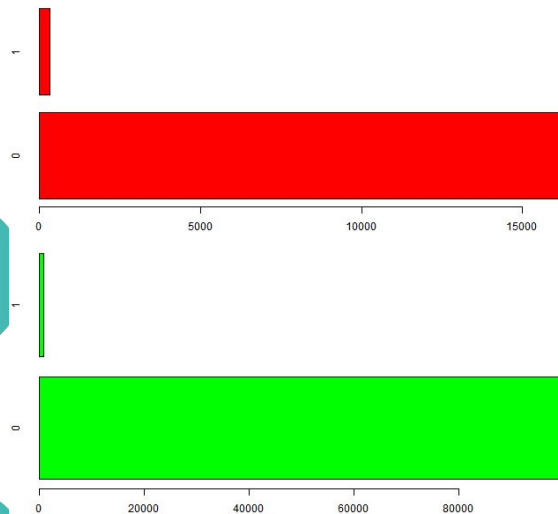
Оценка качества

- $AUC = 0.84$
- Cross val score = 0.845

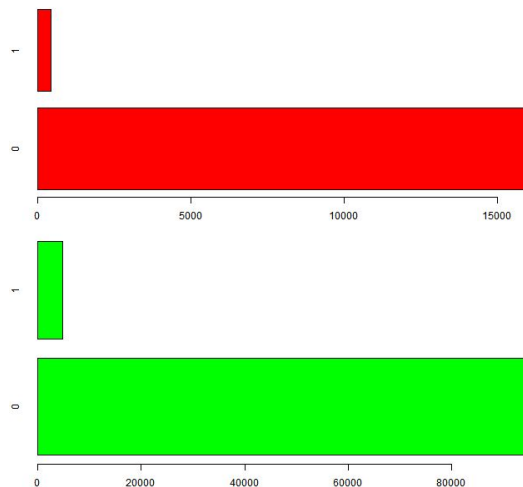


Индивидуальный вклад переменных

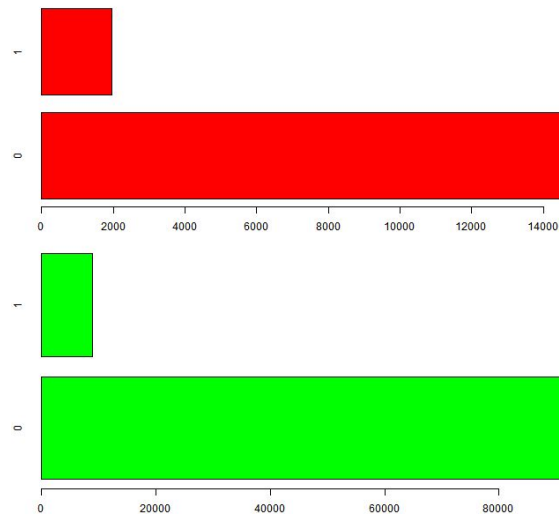
End_TA



PP_Change_Max_Max



HandsetChange_Flag

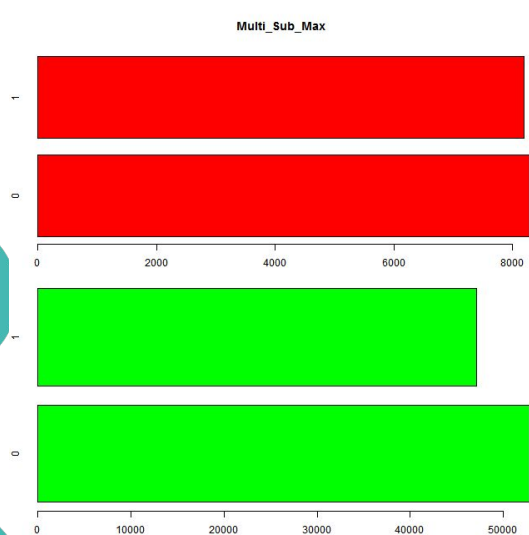


Конец рассрочки на оборудование в последний месяц до контрольного периода

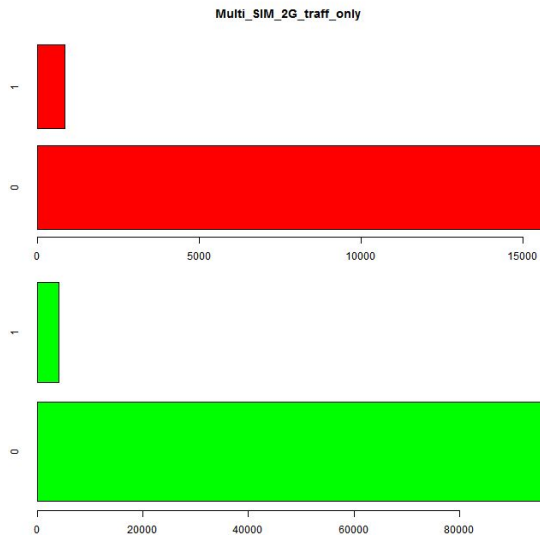
Факт смены тарифного плана в предыдущие 3 месяца

Факт смены устройства за предыдущие 3 месяца

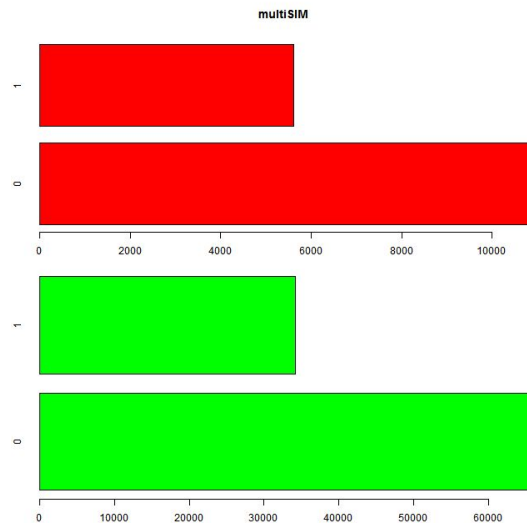
Индивидуальный вклад переменных



Клиент с несколькими сим-картами



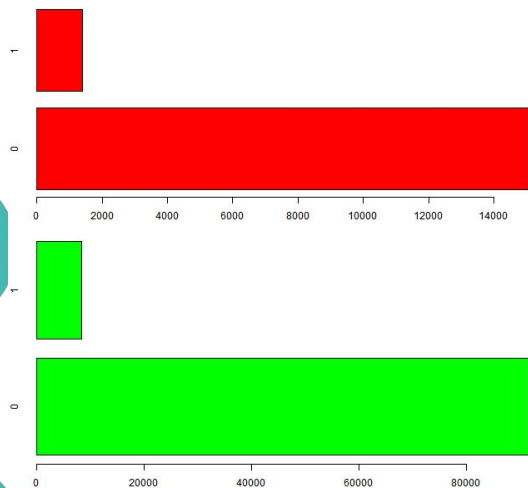
Абонент с мульти-сим устройством
и 2G профилем



Клиент с мульти-сим устройством

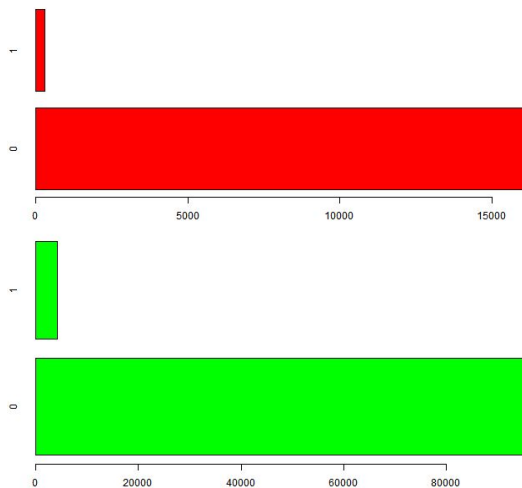
Индивидуальный вклад переменных

shock_data_b



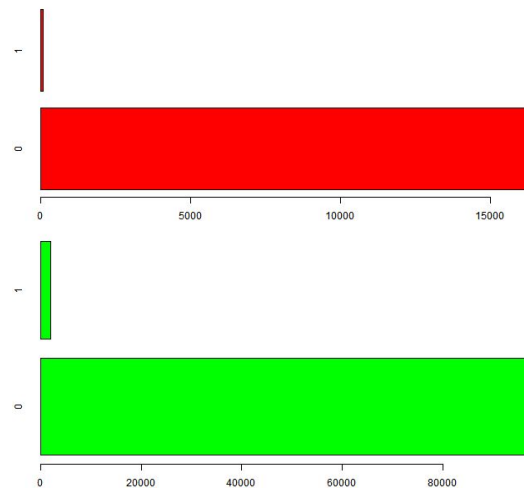
Факт существенного увеличения
счёта за интернет-трафик на
интервале в 3 месяца

shock_int_b



Факт существенного увеличения
счёта за международные звонки на
интервале в 3 месяца

shock_roam_b



Факт существенного увеличения
общего счёта за роуминг на
интервале в 3 месяца

Анализ факторов

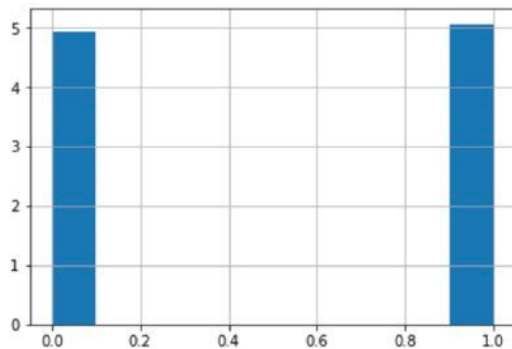
Визуализация feature importance с помощью eli5

Weight	Feature
0.1006	D_ActiveDays
0.0766	D_OutVoice_Velcom_Dur_strange_small
0.0677	D_OutVoice_Velcom_Dur
0.0579	D_ActiveDays_strange_small
0.0470	k3_0
0.0340	Lifetime_Months
0.0280	D_InVoice_Velcom_Dur_strange_small
0.0275	S_BARBL
0.0203	In_Velcom_Uniq
0.0194	Billed_Lifetime
0.0185	D_InVoice_Velcom_Dur
0.0155	shock_data_b
0.0139	k7_3
0.0127	Offnet_Usage_Sum_Sum
0.0127	D_ARPU_strange_small
0.0121	ActiveDays
0.0114	PAYMENTS_AMOUNT
0.0114	R_InVoice_Velcom_Dur
0.0108	OutVoice_Competitor2_Cnt
0.0107	Total_Fee

Multi_Sim | Клиент с мульти-сим устройством

```
train[(train['multiSIM'] == 0) &  
      (train['Lifetime_Months'] > 0.005) &  
      (train['X3G_Data_Usage_Sum_Sum'] < 0.7) &  
      (train['X3G_Data_Usage_Sum_Sum'] > 0.3) &  
      (train['In_Velcom_Uniq'] < 0.31) &  
      (train['Lifetime_Months'] < 0.02)][ 'churn'].hist(normed = True)
```

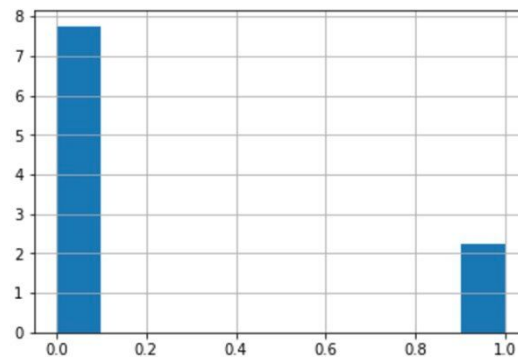
<matplotlib.axes._subplots.AxesSubplot at 0x12bae3c50>



Размер выборки около 500
абонентов

```
train[(train['multiSIM'] == 1) &  
      (train['Lifetime_Months'] > 0.005) &  
      (train['X3G_Data_Usage_Sum_Sum'] < 0.7) &  
      (train['X3G_Data_Usage_Sum_Sum'] > 0.3) &  
      (train['In_Velcom_Uniq'] < 0.31) &  
      (train['Lifetime_Months'] < 0.02)][ 'churn'].hist(normed = True)
```

<matplotlib.axes._subplots.AxesSubplot at 0x12bb7ce50>

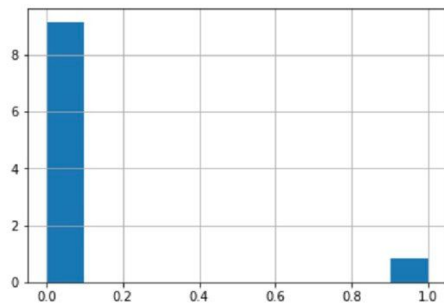


Размер выборки около 800
абонентов

End_TA | Конец рассрочки

```
train[
    (train['End_TA'] == 1) &
    (train['Lifetime_Months'] < 0.035)
]['churn'].hist(normed = True)
```

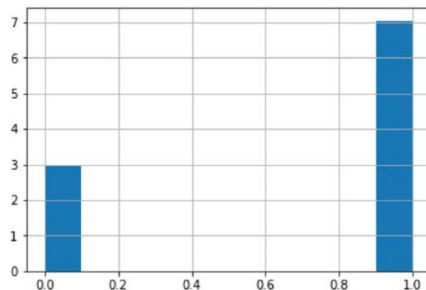
<matplotlib.axes._subplots.AxesSubplot at 0x12cae9150>



Размер выборки около 120
абонентов

```
train[
    (train['End_TA'] == 1) &
    (train['Lifetime_Months'] > 0.035) &
    (train['Lifetime_Months'] < 0.045)
]['churn'].hist(normed = True)
```

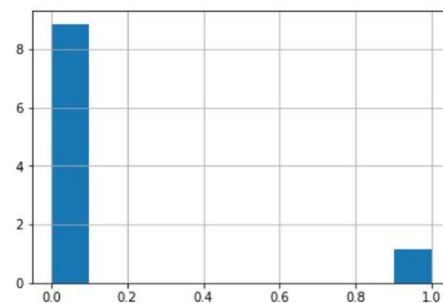
<matplotlib.axes._subplots.AxesSubplot at 0x12d953550>



Размер выборки около 350
абонентов

```
train[
    (train['End_TA'] == 1) &
    (train['Lifetime_Months'] > 0.045)
]['churn'].hist(normed = True)
```

<matplotlib.axes._subplots.AxesSubplot at 0x12cd79850>

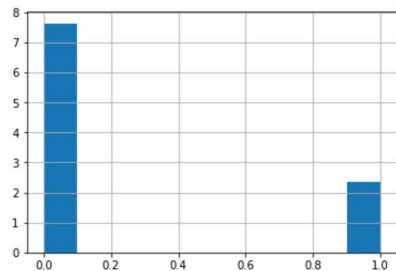


Размер выборки около 600
абонентов

Конец рассрочки / Исходящие на другие сети

```
train[  
    (train['OutVoice_Competitor2_Cnt'] > 0.6)  
][['churn']].hist(normed = True)
```

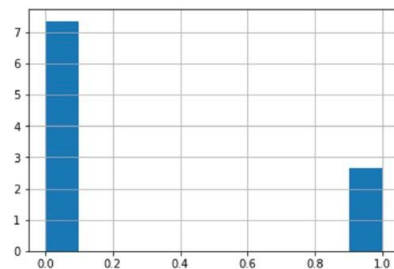
<matplotlib.axes._subplots.AxesSubplot at 0x12f3ce1d0>



Размер выборки около 120
абонентов

```
train[  
    (train['shock_data_b'] == 1) &  
    (train['OutVoice_Competitor2_Cnt'] > 0.6)  
][['churn']].hist(normed = True)
```

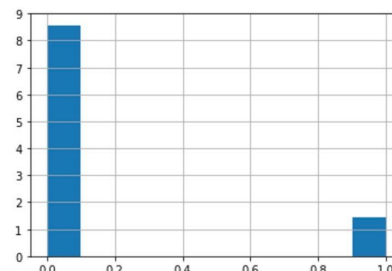
<matplotlib.axes._subplots.AxesSubplot at 0x12ee81150>



Размер выборки около 350
абонентов

```
train[  
    (train['shock_data_b'] == 1)  
][['churn']].hist(normed = True)
```

<matplotlib.axes._subplots.AxesSubplot at 0x12f122f10>



Размер выборки около 600
абонентов

Концепция CRM

Дополнительные устройства

Акция на конец рассрочки

Нотификация о превышении трафика

Предложение смены тарифа с большим
объемом данных

Предложение смены тарифа
ориентированного на другие сети

kb team



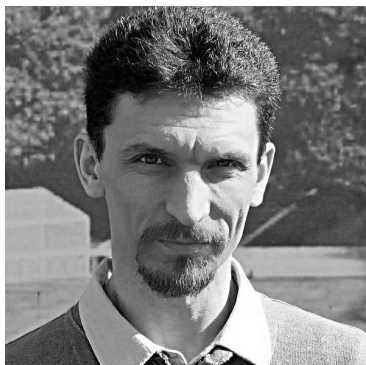
Михаил Койпиш



Александр Короленя



Михаил Мицкевич



Дмитрий Тарасов



Денис Корабков



Ольга Бушуева