



# **Принцип составления индивидуальных предложений абонентам Мегафон**

Савельев Глеб

Факультет Искусственного интеллекта GeekBrains

# Информация о модели

Оценщик:

- CatBoostClassifier 

Гиперпараметры:

- n\_estimators: 1500
- max\_depth: 5
- cat\_features: передал список категориальных признаков

# Особенности

- Модель обучена на сбалансированном датасете

Из-за того, что размер тренировочного датасета был урезан, финальная метрика оказалась ниже, чем у baseline-модели, но такими жертвами удалось добиться бОльшей бизнес-ценности прогнозов и увеличить recall на 93%.

- Применён отбор признаков

В ходе исследования данных обнаружены признаки, которые имеют нулевое значение “важности” (значения получены встроенным в CatBoost методом `get_feature_importance`). Таким образом удалось ужать датасет в ширину практически в 2 раза.

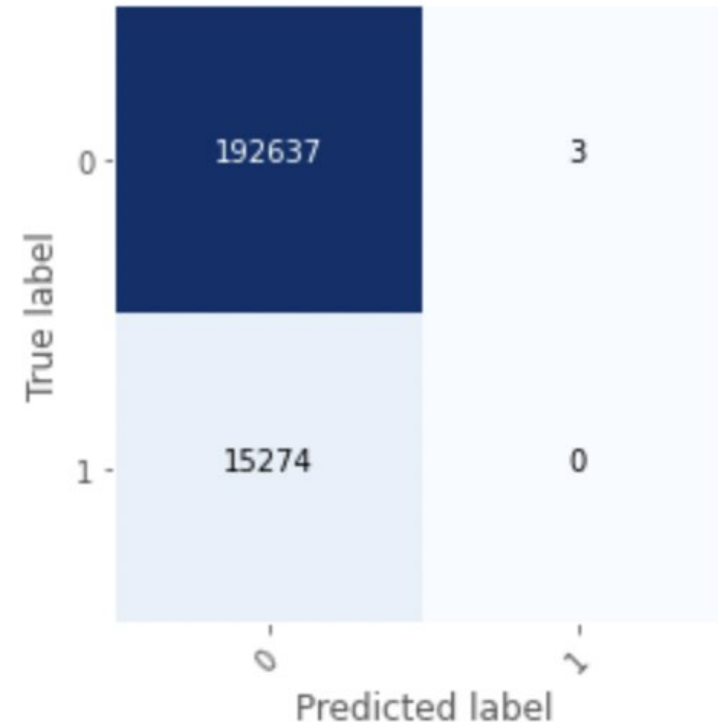
# Результаты

## Baseline-модель

+ относительно хорошее общее качество

– плохо выполняет бизнес-задачу

Baseline Random Forest: confusion matrix



# Результаты

## Финальная модель

- не лучшее общее качество
- + намного лучше выполняет бизнес-задачу
- + быстрее работает за счёт “сужения” датасета

Catboost NearMiss Model: confusion matrix

True label	0	1
	68075	124565
0	1036	14238
1		
Predicted label		

# Обоснование

## Случайный лес

	precision	recall	f1-score	support
0.0	0.99	0.11	0.19	192640
1.0	0.08	0.98	0.15	15274
accuracy			0.17	207914
macro avg	0.53	0.55	0.17	207914
weighted avg	0.92	0.17	0.19	207914

## Градиентный бустинг

	precision	recall	f1-score	support
0.0	0.99	0.35	0.52	192640
1.0	0.10	0.93	0.18	15274
accuracy			0.40	207914
macro avg	0.54	0.64	0.35	207914
weighted avg	0.92	0.40	0.50	207914

# Принцип составления индивидуальных предложений для выбранных абонентов

Для составления индивидуального предложения клиенту модель, зная его профиль покупателя за предыдущий период, сможет определить, с какой вероятностью клиент приобретёт услугу.

Далее в зависимости от бизнес-цели - охватить больше возможно заинтересованных в покупке или найти заинтересованных как можно точнее, выставляется порог вероятности, после которого клиент будет считаться заинтересованным в покупке. Если модель сообщает о том, что вероятность приобрести услугу клиентом выше порога, то такому клиенту следует порекомендовать соответствующую услугу.