**Цель проекта:**

* Необходимо построить модель прогнозирования ухода клиентов из банка на основе исторических данных о поведении клиентов и расторжении договоров с банком.
* Качественной будет считаться модель с метрикой F1 не менее 0.59.

**Тестируемые модели:**

Были обучены и оценены 3 модели:

* Дерево решений
* Модель леса
* Логистическая регрессия

Ниже представлены результаты моделей

f1 results

|  | **basic** | **upsampling** | **balancing** | **downsampling** |
| --- | --- | --- | --- | --- |
| **dtc** | 0.557427 | 0.555556 | 0.555556 | 0.553191 |
| **rfc** | 0.568966 | 0.604651 | 0.607595 | 0.593074 |
| **lr** | 0.300412 | 0.495661 | 0.495069 | 0.500000 |

auc-roc results

|  | **basic** | **upsampling** | **balancing** | **downsampling** |
| --- | --- | --- | --- | --- |
| **dtc** | 0.806275 | 0.815042 | 0.815016 | 0.809483 |
| **rfc** | 0.836066 | 0.839837 | 0.848921 | 0.849277 |
| **lr** | 0.772518 | 0.774835 | 0.774782 | 0.775518 |

**Приоритетной является модель RandomFores**t, полученная с учетом балансировки классов. Модель показала высокий уровень метрик качества на валидационной выборке:

* F1 = 0.61 (превышение целевого показателя)
* AUC-ROC = 0.84 (топ-2 результат)

**На тестовой выборке F1 = 0.60 (превышение целевого показателя)**

**Параметры модели:**

* Кол-во оценщиков: 90
* Глубина: 10
* Порог классификации: 0.5
* class\_weight='balanced'