**Исследование показателей комплекса стенда SWaT, анализ данных показателей стенда в стационарном состоянии и при воздействии атак и создание оптимальной модели предсказания состояния с помощью моделей бинарной классификации**

**Данные**

В наличии были ориентировочные данные о процессе по следующим этапам процесса очистки воды на полигоне без указания размерностей и прочих параметров данных:

* P1: забор сырой воды
* P2: обеззараживание химикатами
* P3: ультрафильтрация
* P4: дехлорирование с помощью ультрафиолетовых ламп
* P5: очистка методом обратного осмоса
* P6: обратная промывка и очистка мембраны ультрафильтрации

**Задача**

В рамках задачи проекта предстоит построить модель, которая предскажет состояние(атака(1), отсутствие атаки(0)).

**Результаты проекта**

Были построены и протестированы различные модели бинарной классификации. Итоговая модель, которая была выбрана лучшей по показателю метрики recall(полнота) стала модель на основе алгоритма LGBMClassifier и ее показатель полноты составляет 0,999740, результирующими признаками данной модели стали - fit\_501, pit\_502, lit\_301, ait\_201, ait\_501, остальные признаки оказались наименее значимыми.

**Используемые библиотеки**

*pandas*, *numpy*, *warnings*, *matplotlib*, *seaborn*, *shap*, *lightgbm*, *sklearn, plotly, imblearn*