

Оптимизация в промышленности

стажировка GlowByte

Литовченко Максим

15 июня 2024 г.

EDA

Были обнаружены особенности в данных:

1. Существуют дубликаты данных и аномалии, которые, судя по всему, несут в себе какую-то информацию
2. Распределения данных имеют несколько мод
3. Большое число выбросов

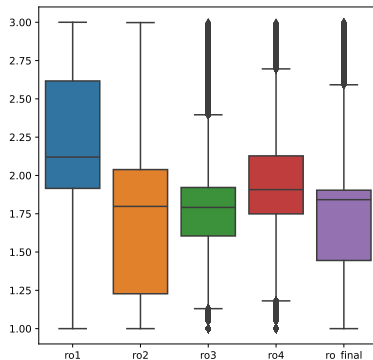


Рис. 1. Распределение ro

2 этап

Было проделано:

1. Замена аномальных данных при помощи KNN
2. Удаление дубликатов с созданием признака на их наличие, для сохранения информации
3. Определено решение через RL Q-Learning

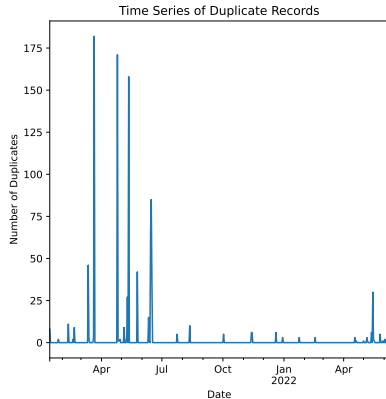


Рис. 2. Распределение дубликатов

3 этап

Была получена сходящаяся модель и выдвинуты предположения об улучшении

The End