

Лабораторная №1. Настройка гиперпараметров и выбор алгоритма

Набор данных

- Выберите любой набор данных для задачи обучения с учителем.
- Можно найти на сайте <https://www.openml.org/search?type=data>
- Не стоит брать слишком большие наборы данных.
- Предобработайте набор данных:
 - нормализуйте;
 - заполните пропуски;
 - приведите признаки к числовому типу.
- Разбейте набор данных на тренировочную, валидационную и тестовую части.

Задание

- Сведите задачу выбора алгоритма к задаче настройки гиперпараметров. Можно использовать любую функцию ошибки.
- Портфолио алгоритмов должно содержать не менее пяти алгоритмов включая:
 - метод ближайших соседей;
 - метод опорных векторов;
 - байесовский классификатор;
 - дерево принятия решений.
- Решите полученную задачу оптимизации при помощи библиотеки Optuna или её аналогов.

Анализ

- Сравните полученный результат со случайным поиском.
- Оцените на тестовой части конфигурации гиперпараметров, которые были найдены разными методами.
- Постройте графики зависимости функции ошибки от номера итерации. Если вы использовали Optuna, выведите также информацию о полезности гиперпараметров.