

Непараметрическая регрессия

Набор данных

Выберите любой понравившийся набор данных для задачи классификации на сайте openml.org. Выбранный набор данных должен содержать не менее 100 объектов и не менее трёх классов. Запрещено брать набор данных **iris**! Не забудьте векторизовать признаки вашего набора данных (перейти от категорий к числам), заполнить пропуски (если есть) и нормализовать.

Сведение к задаче регрессии

Перейдите от задачи классификации к задаче регрессии двумя разными способами:

1. Наивный способ. Каждое значение класса представляется одним числом. Во время предсказания полученный ответ округляется до ближайшего целого числа.
2. Используя [OneHot преобразование](#). Вместо одного целевого признака в набор данных добавляется столько новых числовых переменных, сколько в нём содержится классов.

Настройка гиперпараметров

Для каждого из преобразований найдите лучшую комбинацию функции расстояния, окна и ядра для метода ближайших соседей. Для лучшего преобразования и найденной комбинации постройте графики зависимости F-меры от числа ближайших соседей или ширины окна. Используйте Leave-One-Out перекрёстную проверку для подсчёта F-меры.