

X­­(1) – самый близкий к новому объект, X*(l)* – самый дальний

Самый популярный класс – который встречается чаще всего

[y(i) = y] – айверсоны (Нотация Айверсона, или скобка Айверсона), внутри них логическое выражение

Если выражение верное, то айверсоны выдают 1, если нет – 0

Внутри функции a(x) мы считаем, сколько объектов из среди k ближайших, относятся к классу y

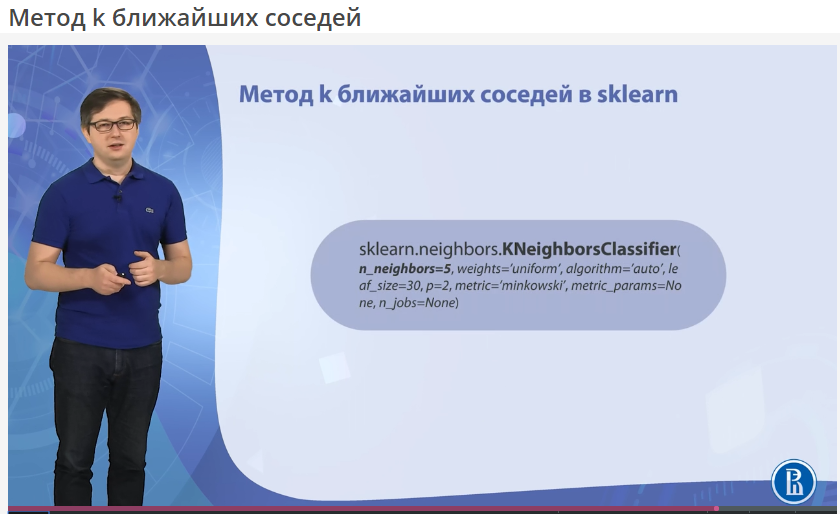


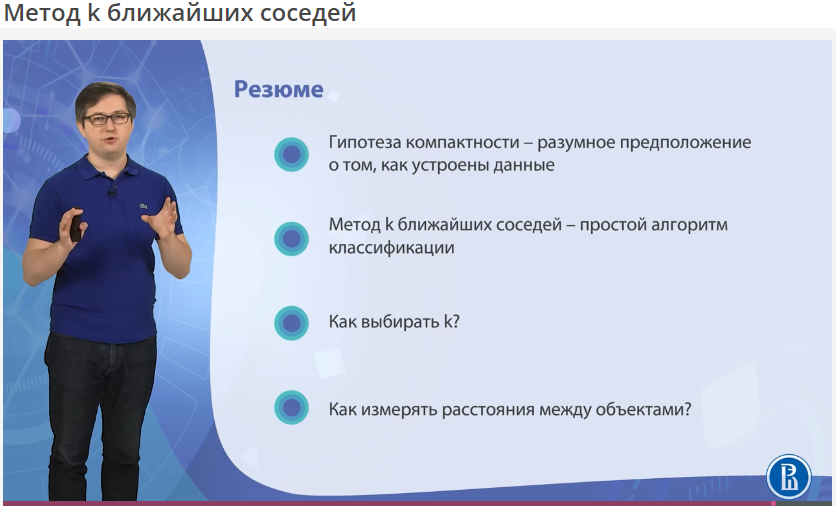
Допустим, объект, отмеченный стрелкой, является новым, и нам надо понять, к какому классу его отнести – к зеленым или к красным.

Если k=3, то ближайшими будут 2 красных и 1 зеленый объект. Тогда новый объект мы отнесем к красному классу

Если k=5, то ближайщими будут 2 красных и 3 зеленых, тогда объект отнесем к зеленому классу.

Кажется, все зависит от того, чему равно k





Давайте проверим, насколько хорошо вы поняли материал прослушанной лекции.

3 из 3 баллов (не оценивается)

1) Мы хотим предсказывать зарплату человека на основе его роста. В обучающей выборке есть 5 человек с таким же ростом, как у вас, и зарплата у всех них примерно равна 50 тысячам рублей. Какую зарплату нужно предсказать для вас, если пользоваться гипотезой компактности?

Гипотеза компактности применима только в задачах классификации

Если в выборке нет других людей с похожим ростом, то 50 тысяч рублей верно

В любом случае 50 тысяч рублей, поскольку достаточно взять часть людей из обучающей выборки с нужным ростом

2) В чём заключается обучение метода k ближайших соседей?

Запоминается обучающая выборка верно

Вычисляются и запоминаются попарные расстояния между всеми объектами обучающей выборки

Выявляются и запоминаются самые характерные объекты в обучающей выборке

Подбирается и запоминается значение k

3) В чём заключается применение метода k ближайших соседей к новому объекту x?

Выбирается k объектов обучающей выборки, расстояние от которых до x не превышает N (это ещё один параметр метода). Выдаётся тот класс, представителей которого больше всего среди этих k объектов

Объекты обучающей выборки сортируются по расстоянию до x, выбираются k ближайших. Выдаётся тот класс, представителей которого меньше всего среди этих k объектов

Выбирается k объектов обучающей выборки, расстояние от которых до x не превышает N (это ещё один параметр метода). Выдаётся тот класс, представителей которого меньше всего среди этих k объектов

Объекты обучающей выборки сортируются по расстоянию до x, выбираются k ближайших. Выдаётся тот класс, представителей которого больше всего среди этих k объектов верно