

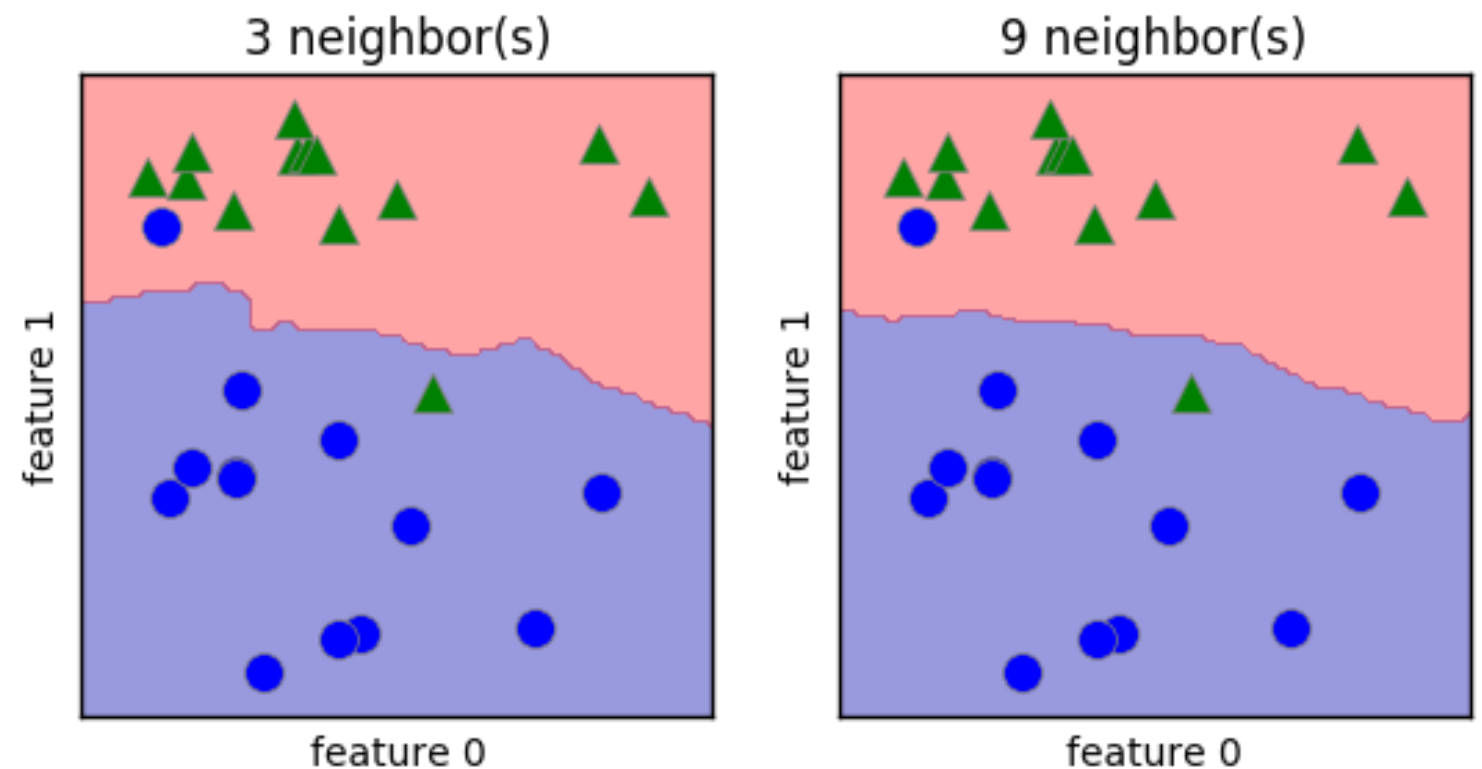
# Метрические методы классификации (kNN)

Довженко Николай  
БПМИ162

# Что такое метрическая классификация?

- **Метрический классификатор** (similarity-based classifier) — алгоритм классификации, основанный на вычислении оценок сходства между объектами.
- Функция сходства  $p(x, y)$  — не обязательно метрика.
- Метрические классификаторы принимают гипотезу о компактности

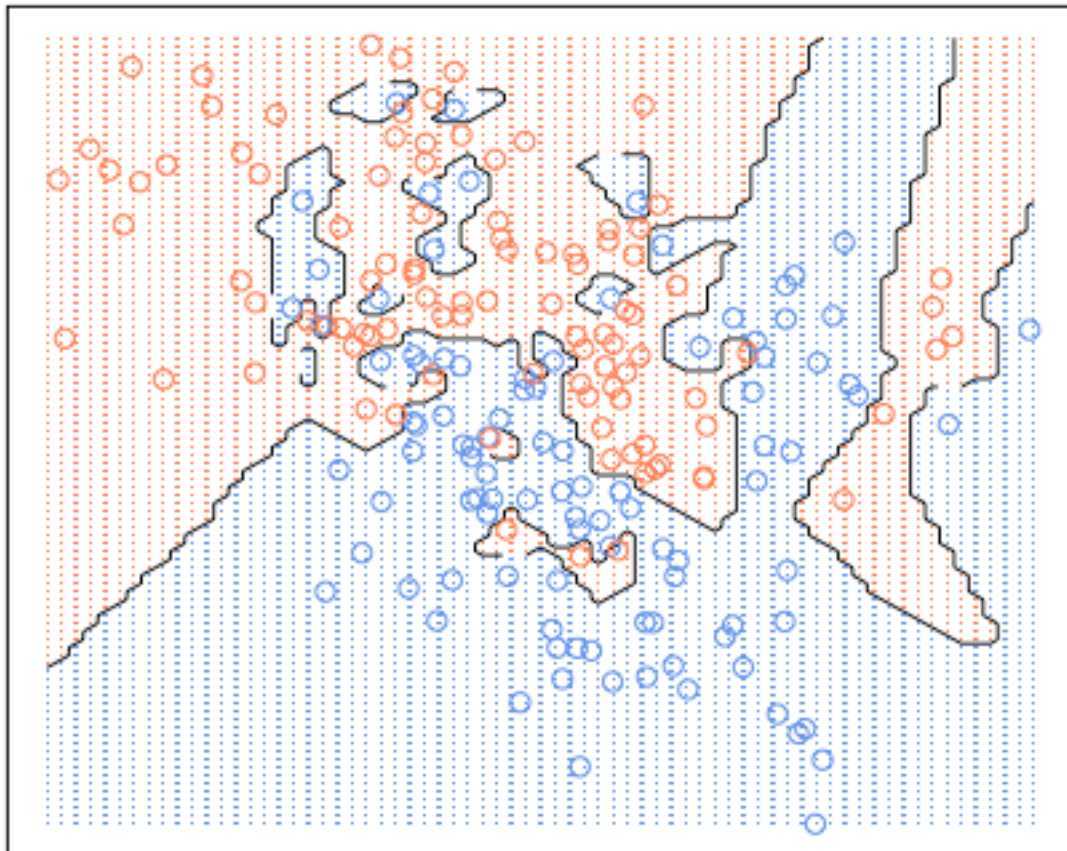
# KNN



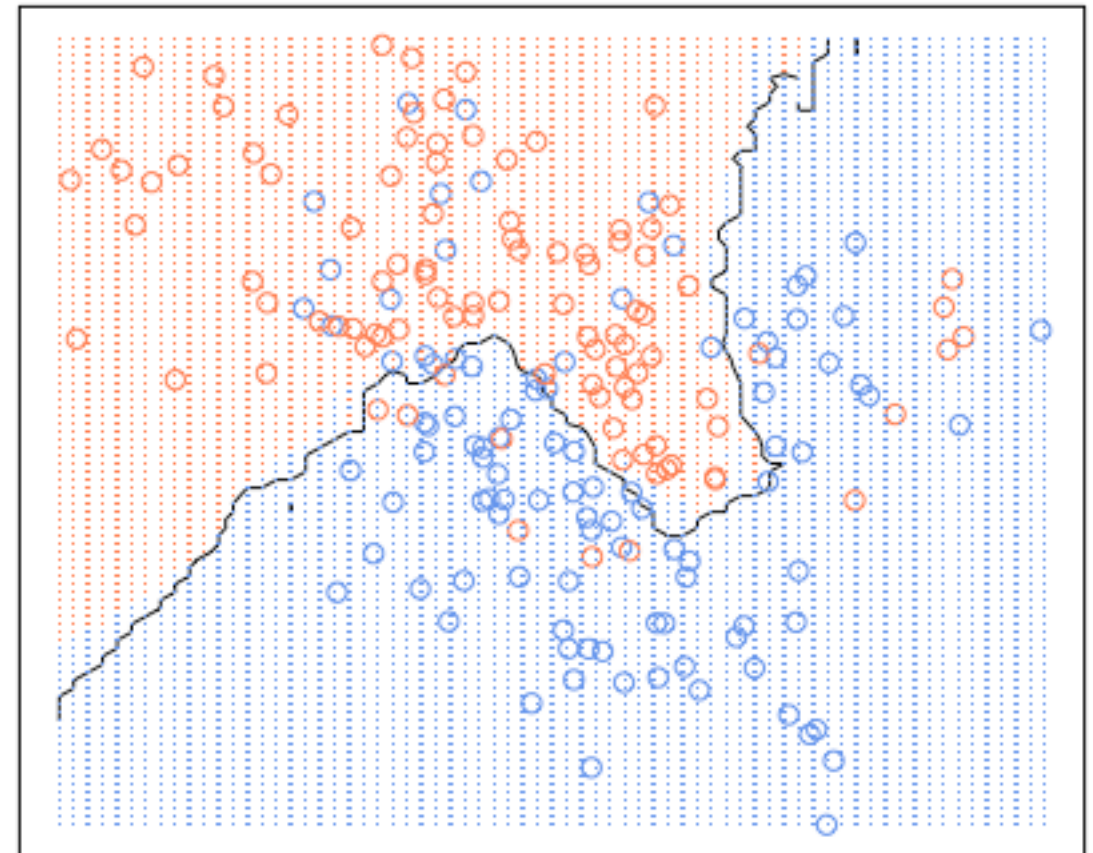
- Метод k ближайших соседей (k nearest neighbors) - простейший алгоритм метрической классификации
- Есть размеченная выборка (обучающая) и неразмеченная (тестовая)
- Класс объекта тестовой выборки определяется через классы k ближайших соседей из тестовой выборки

# Тюнинг гиперпараметров

nearest neighbour ( $k = 1$ )



20-nearest neighbour



# Вероятность ошибки

- Метод ближайшего соседа очень точный, если размер обучающей выборке стремится к бесконечности
- Метод ближайших соседей был оценен в 1976 году

$$R^* \leq R_{k\text{NN}} \leq R^* \left( 2 - \frac{MR^*}{M-1} \right)$$

# Проблемы

- У классов различная плотность распределения
- Слишком высокая размерность пространства объектов
- Слишком большой размер выборки

# Другие метрические алгоритмы классификации

- Метод потенциальных функций
- Метод радиальных базисных функций
- Метод парзеновского окна
- Метод дробящихся эталонов
- Алгоритм вычисления оценок

**Спасибо за  
внимание!!**