

## Лабораторная работа # 4

### Метрические методы классификации.

Цель лабораторной работы: исследовать качество алгоритма классификации в зависимости от характера метрики и других параметров метрического классификатора.

#### Постановка задачи

1. Разбить данные на обучающую и тестовую выборки (например, в соотношении 80/20).
2. Обучить алгоритм  $k$  ближайших соседей из библиотеки `scikit-learn` (`KNeighborsClassifier`) при различных наборах параметров:
  - Количество соседей  $n\_neighbors$ : не менее 5 различных значений. Обязательно рассмотреть случай  $n\_neighbors = 1$ ;
  - Вес объектов  $weights$ : рассмотреть 'uniform' и 'distance';
  - Метрика: метрика городских кварталов ('cityblock'), евклидова метрика ('euclidean') и метрика Чебышева (`scipy.spatial.distance.chebyshev`).
3. Провести сравнительную оценку качества моделей на основе тестовой выборки. В качестве метрики классификации выбрать любую знакомую метрику (accuracy, precision&recall,  $F_1$  и т.д.).
4. Определить модель с наилучшей комбинацией параметров.

#### Рекомендации

- Для разбиения выборки можно использовать функцию `"train_test_split"`;
- Для проведения сравнительной оценки качества моделей рекомендуется использовать метод `GridSearchCV`.