

# Ирисы

## Цель

ознакомиться с набором данных "Ирисы" и его характеристиками, научиться работать с пропущенными и аномальными данными, обучаться модели дерева решений и анализировать влияние параметров на её качество.

### Что нужно сделать:

- O1** Изучите набор данных "Ирисы", который включает в себя три вида цветов ириса и различные характеристики каждого цветка.
- O2** Выполните следующие шаги:
  - a)** Загрузите набор данных "Ирисы" из библиотеки scikit-learn и проверьте его на наличие пропусков и аномалий.
  - b)** Создайте и обучите модель дерева решений, используя 80% данных для тренировки и 20% для тестирования.
  - c)** Получите метрики производительности вашей модели (точность, полнота, F1-мера) и представьте результаты в таблице.
  - d)** Проанализируйте, как различные параметры (например, количество минимальных образцов для разбиения) влияют на качество модели. Постройте график зависимости точности от глубины дерева.

### Критерии оценивания

- |           |   |        |
|-----------|---|--------|
| <b>K1</b> | Понимание характеристик набора данных и его структуры.  | 1 балл |
| <b>K2</b> | Проверка набора данных на наличие пропущенных значений и аномалий, их корректная обработка.   | 1 балл |
| <b>K3</b> | Создание и успешное обучение модели дерева решений, использование 80% данных для тренировки и 20% для тестирования.   | 1 балл |
| <b>K4</b> | Получение метрик (точность, полнота, F1-мера) и представление их в таблице.   | 1 балл |
| <b>K5</b> | Анализ влияния различных параметров (например, количество минимальных образцов для разбиения) на качество модели и построение графика зависимости точности от глубины дерева. | 1 балл |

**Максимальное** количество баллов **5 баллов**

**Минимальное** количество баллов чтобы преподаватель смог зачесть вашу работу **2 балла**