

Задача В. «Идентификация продукта»



Умная камера фотосепаратора (аппарата для сортировки сыпучих продуктов) настроена на выявление признаков подсолнечника, ячменя, пшеницы и гороха.

В файле «Train.csv» представлены признаки 4-х видов сыпучих продуктов:

- **Area** – площадь объекта;
- **MajorAxisLenght** – длина главной оси инерции;
- **MinorAxisLenght** – длина побочной оси инерции;
- **Eccentricity** – эксцентриситет объекта;
- **Orientation** – угол ориентации объекта;
- **ConvexArea** – площадь выпуклой оболочки, построенной над объектом;
- **FilledArea** – заполненность;
- **EulerNumber** – число Эйлера;
- **EquivDiameter** – эквивалентный диаметр;
- **Solidity** – твердость;
- **Extent**;
- **Perimeter** – периметр объекта;

Целевая переменная **Class** – класс объекта, принимает 4 значения.

Ваша задача состоит в том, чтобы по имеющимся признакам классифицировать продукты по видам: ячмень, подсолнечник, пшеница и горох.

Обращаем ваше внимание: в разметке могут быть ошибки, так как продукт содержит примеси. Это важно учитывать при анализе данных и построении классификационной модели.

Входные данные: Файл «Train.csv» содержит признаки объектов и их маркировку.

Выходные данные: для каждой строки файла «Test.csv» класс соответствия объекту. Ответ представить в виде txt файла.

Оценка качества модели: average f1 score. Подробнее по ссылке: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html