

Машинное обучение.

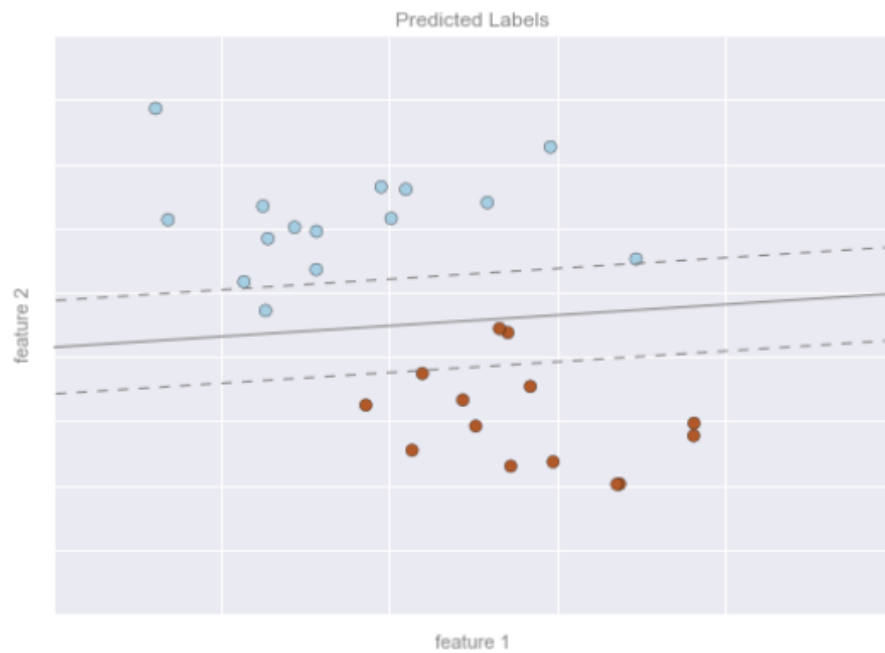
Категории машинного обучения.

На базовом уровне машинное обучение можно разделить на два основных типа:

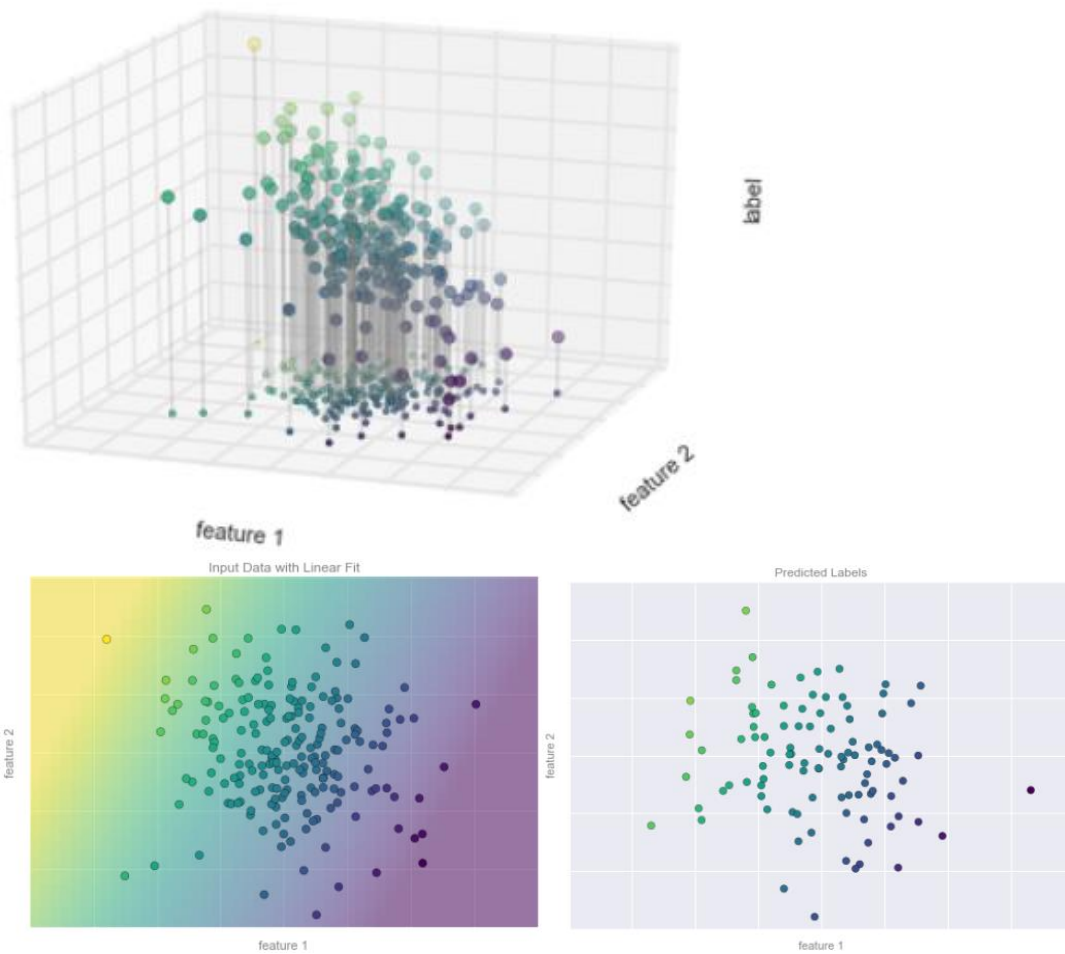
- **Машинное обучение с учителем** - включает моделирование признаков данных и соответствующих данным меток. Оно разделяется далее на **задачи классификации** (метки представляют собой дискретные категории) и **задачи регрессии** (метки являются непрерывными величинами).
- **Машинное обучение без учителя** - включает моделирование признаков набора данных без меток, можно описать фразой "Пусть набор данных говорит сам за себя". Эти модели включают в себя такие задачи, как **кластеризация** (для выделения отдельных групп данных) и **понижение размерности** (поиск более сжатых представлений данных).

* Существует еще метод частичного обучения (наличие лишь неполных меток).

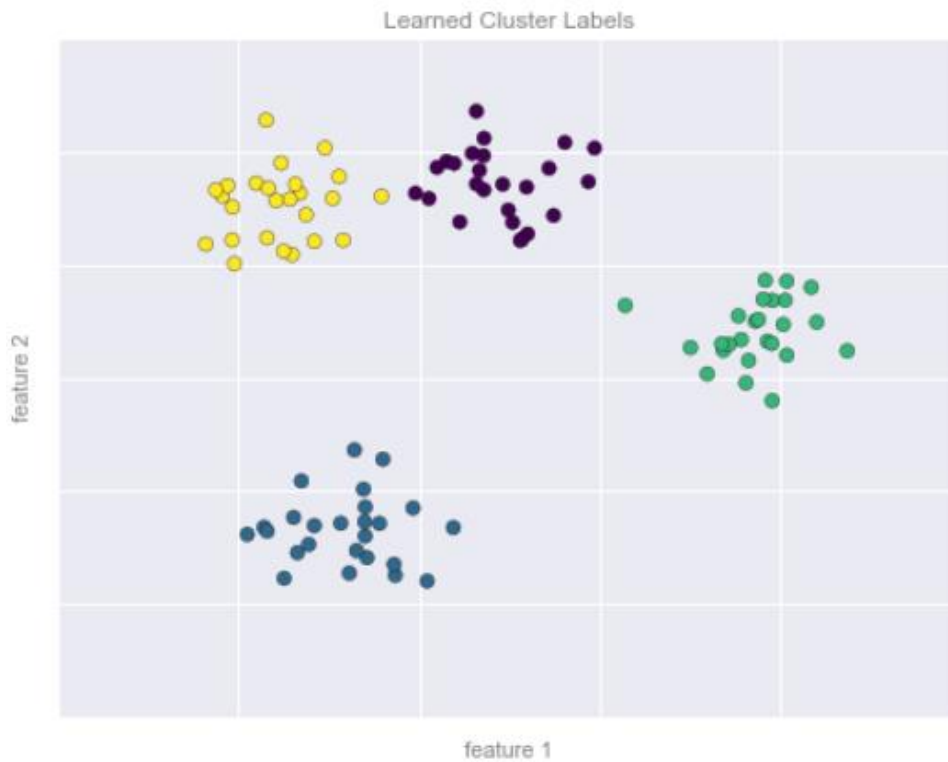
- **Классификация** - модели для предсказания меток из двух или более отдельных категорий;



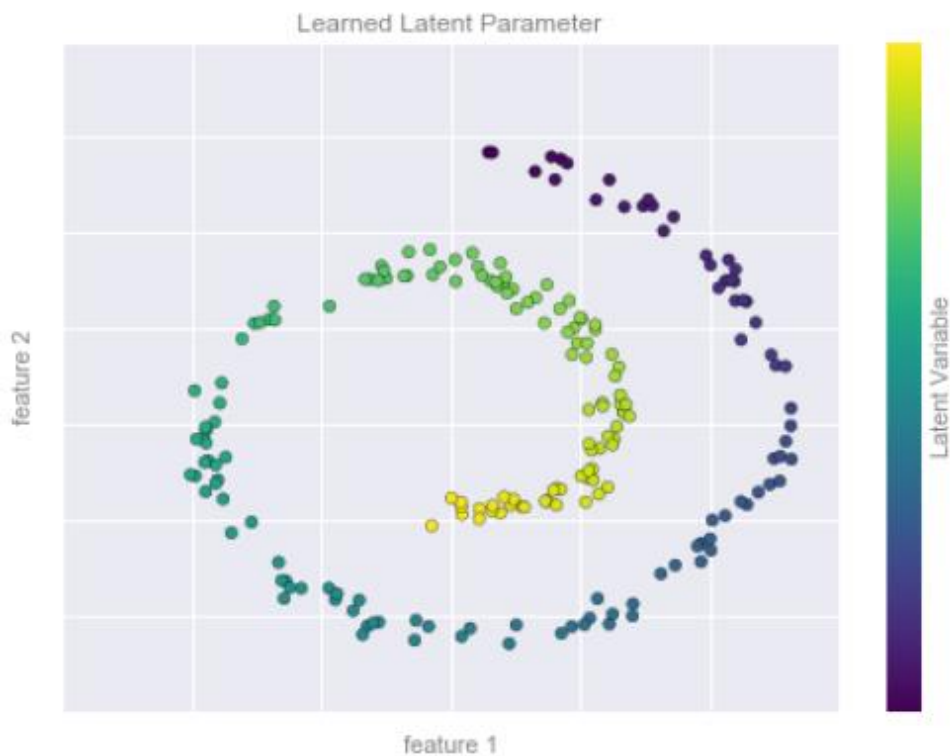
- **Регрессия** - модель для предсказания непрерывных меток;



- **Кластеризация** - модели для выявления и распознавания в данных отдельных групп;



- **Понижение размерности** - модели для выявления и распознавания низкоразмерной структуры в высокоразмерных данных;



API статического оценивания библиотеки Scikit-Learn

Основывается на следующих принципах:

- *Единообразие* - интерфейс всех объектов идентичен
- *Контроль* - видимость всех задаваемых значений параметров
- *Ограниченная иерархия объектов* - массивы NumPy, df Pandas.
- *Объединение* - многие МЛ задачи можно выразить алгоритмами более низкого уровня, библиотека соблюдает эти правила.
- *Разумные значения по умолчанию* - библиотека задает для необходимых моделей пользовательских параметров соответствующие значения по умолчанию.

Основы API статического оценивания (шаги):

1. Выбор класса модели с помощью импорта соответствующего класса оценщика из библиотеки.
2. Выбор гиперпараметров модели путем создания экземпляра этого класса с соответствующими значениями.
3. Компоновка данных в матрицу признаков и целевой вектор в соответствии с описанием выше.
4. Обучение модели на своих данных посредством вызова метода **fit()** экземпляра модели.
5. Применение модели к новым данным:
 - В случае машинного обучения с учителем метки для неизвестных данных обычно предсказывают с помощью метода **predict()**;
 - В случае машинного обучения без учителя выполняется преобразование свойств данных или вывод их значений посредством методов **transform()** или **predict()**.