

Обучение без учителя

Unsupervised Learning

Обучение без учителя (Unsupervised Learning)

- Обучение с учителем (Supervised Learning)
 - При помощи **размеченных (labeled)** данных предсказать значение для новых данных (регрессия или классификация).
- Обучение без учителя (Unsupervised Learning)
 - При помощи **неразмеченных (unlabeled)** данных найти закономерности в данных, кластеры или существенные компоненты.

Обучение без учителя (Unsupervised Learning)

- Обучение без учителя (Unsupervised Learning):
 - Кластеризация:
 - С помощью признаков сгруппировать строки в отдельные кластеры.
 - Уменьшение размерности:
 - С помощью признаков перейти к меньшему количеству компонент в данных.

Обучение без учителя (Unsupervised Learning)

- Смена парадигмы при переходе от обучения с учителем к обучению без учителя:
 - **Метрики для обучения с учителем не подходят для обучения без учителя!**
 - Как мы можем сравниваться с правильными ответами, если у нас нет известных значений целевой переменной?

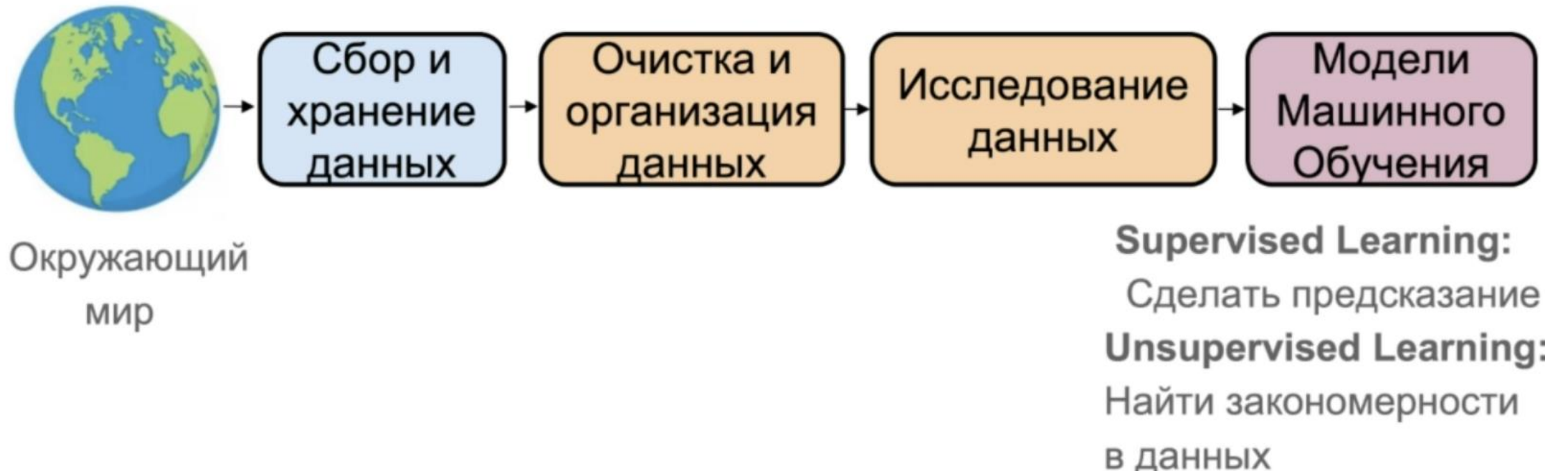
Обучение без учителя (Unsupervised Learning)

- Вместо таких метрик как RMSE и Accuracy, нам нужно как-то по-другому оценивать работу моделей обучения без учителя.
- Даже само понятие “насколько хорошо работает модель” придётся сформулировать по-другому!

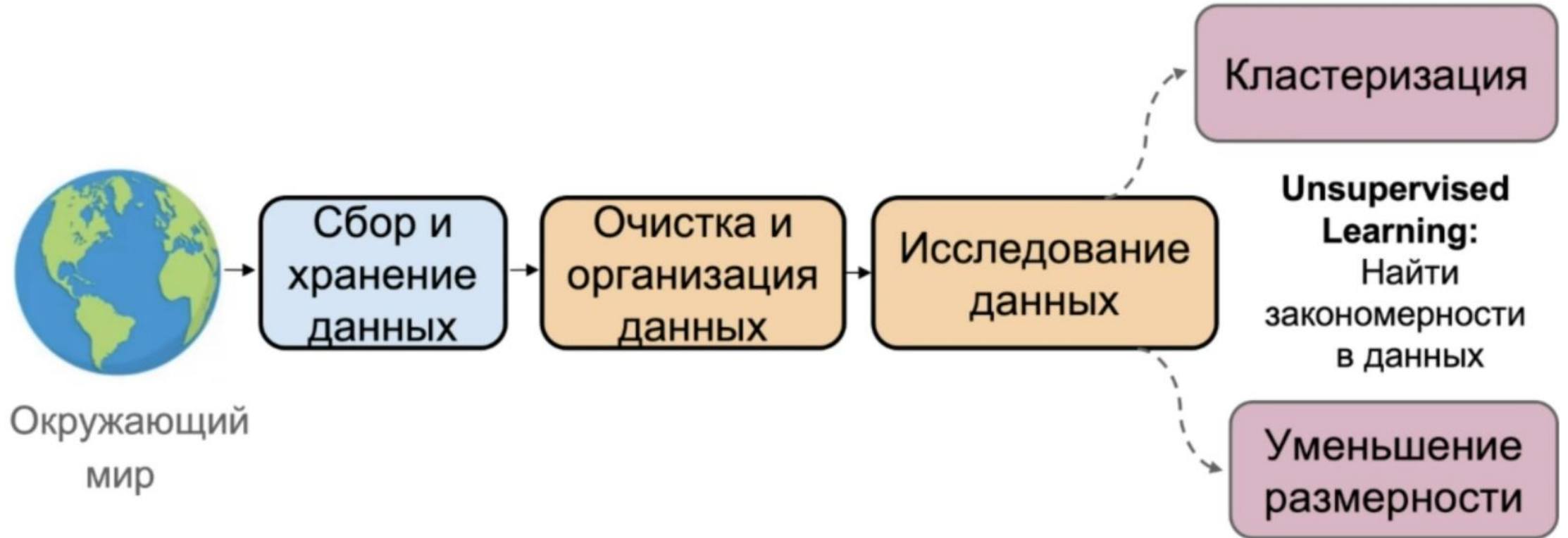
Этапы работ по машинному обучению



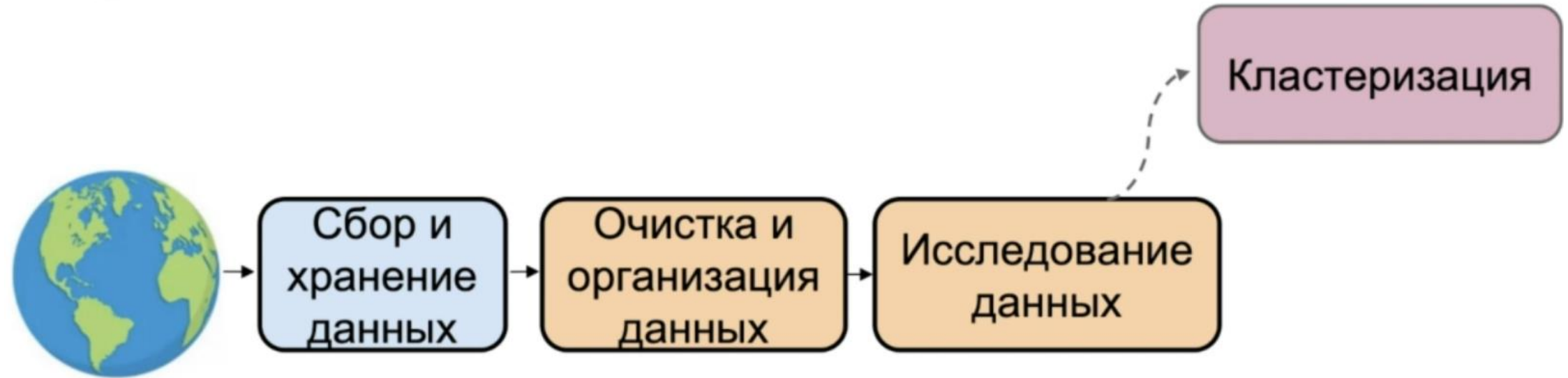
Этапы работ по машинному обучению



Этапы работ по машинному обучению



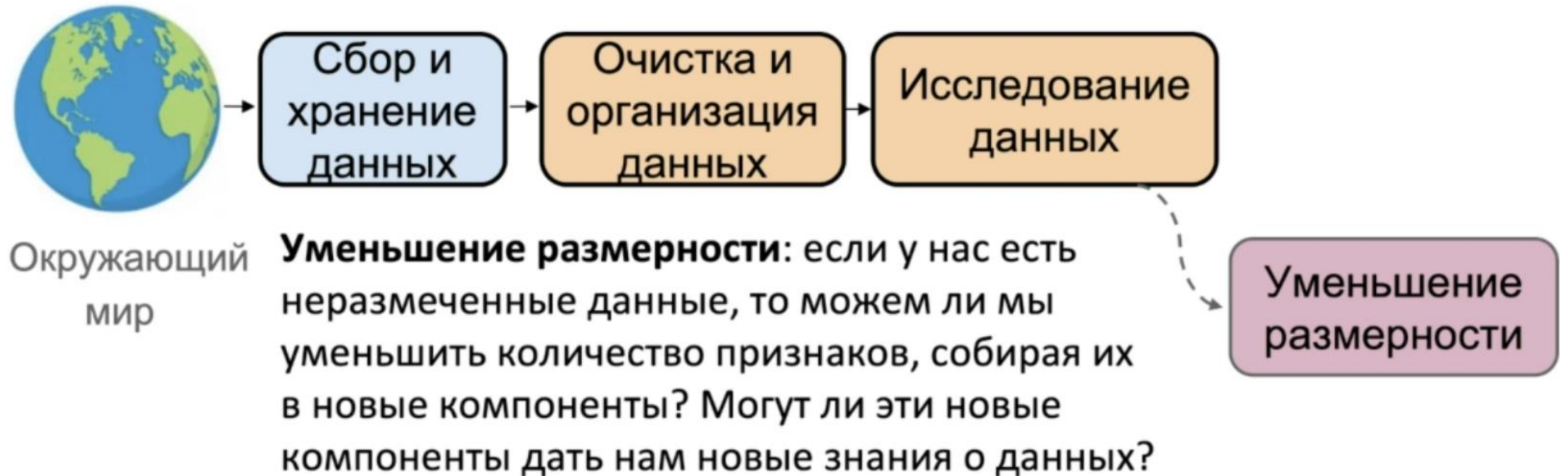
Этапы работ по машинному обучению



Окружающий
мир

Кластеризация: если у нас есть неразмеченные данные, то можем ли мы собрать похожие точки данных вместе, чтобы “найти” возможные значения целевой переменной для кластеров?

Этапы работ по машинному обучению



Обучение без учителя (Unsupervised Learning)

- Мы начнём с изучения методов кластеризации данных, таких как K-Means и Иерархическая кластеризация, а затем перейдём к уменьшению размерности.
- Попутно мы узнаем о методах интерпретации результатов моделей, поскольку оценка работы модели - это более тонкий вопрос для обучения без учителя.

Обучение без учителя (Unsupervised Learning)

- Какие вопросы следует иметь ввиду:
 - Что означает фраза, что мы можем “найти” значения целевой переменной (label) с помощью кластеризации данных?
 - Как мы можем оценить работу модели, не имея значений целевой переменной?
 - Могут ли комбинации признаков содержать полезную информацию о данных?