

ПОНЯТИЕ МАШИННОГО ОБУЧЕНИЯ

Кто занимается анализом данных



Josh Wills

@josh_wills



 Follow

Data Scientist (n.): Person who is better at statistics than any software engineer and better at software engineering than any statistician.

 Reply  Retweet  Favorite  More

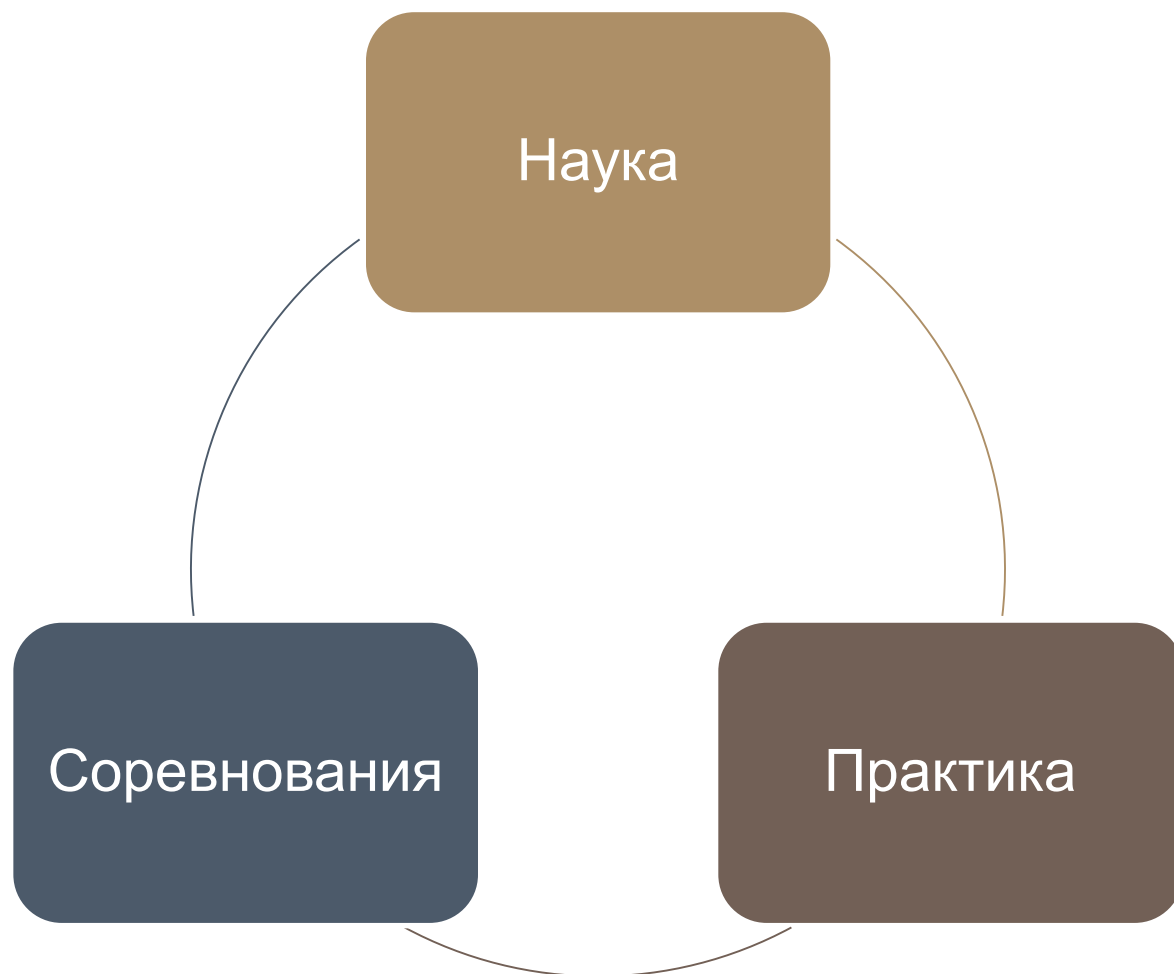
RETWEETS
907

FAVORITES
418

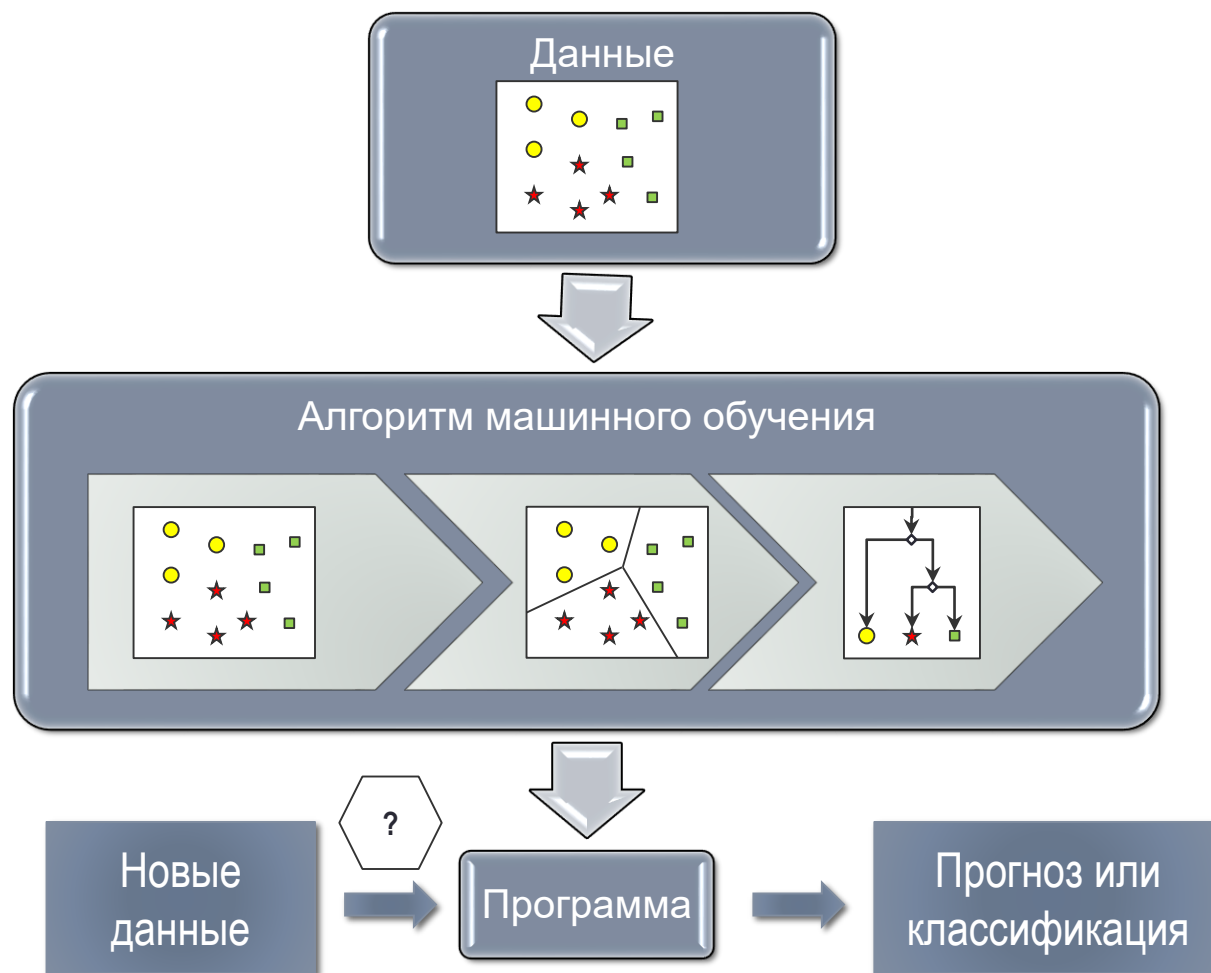


12:55 PM - 3 May 2012

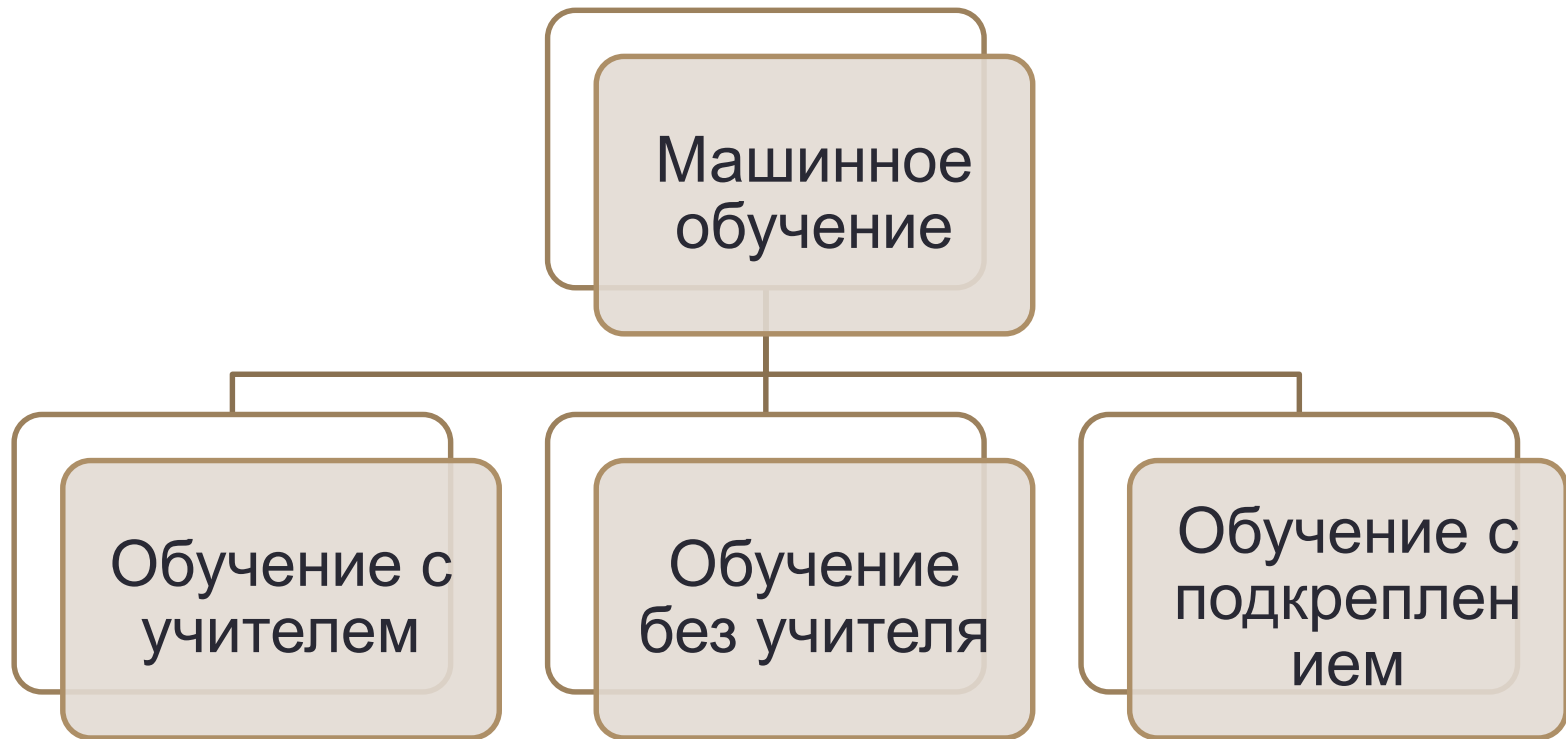
Как занимаются анализом данных



Понятие машинного обучения



Основные виды машинного обучения



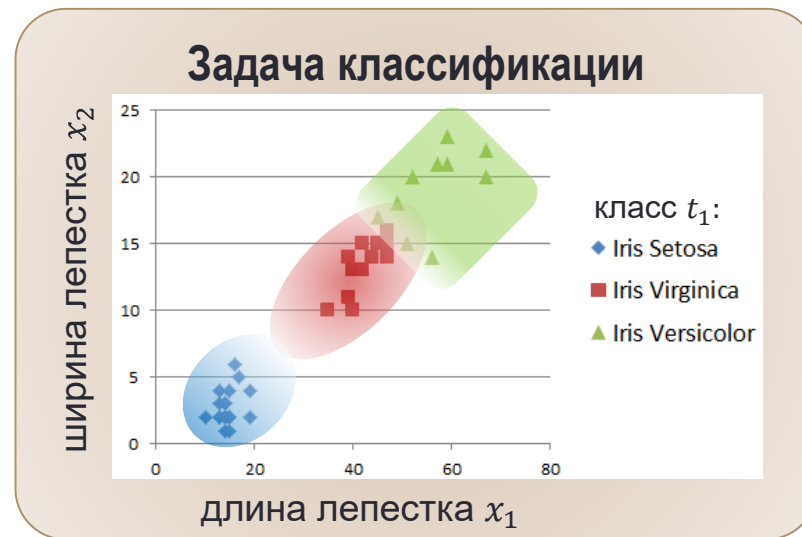
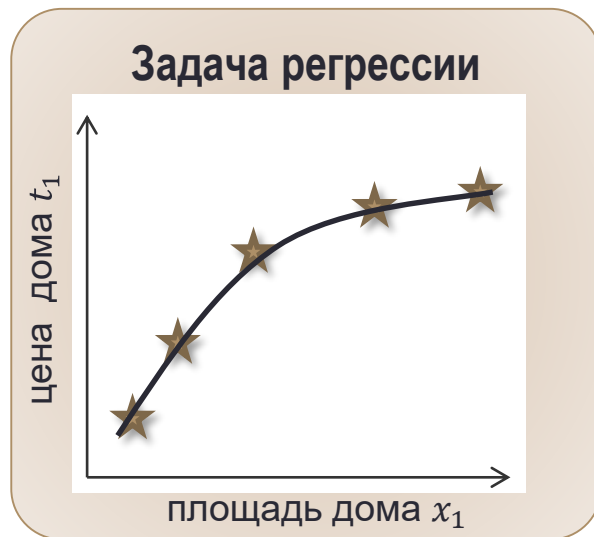
Задачи машинного обучения

- Цель: определить форму функциональной зависимости
 - Независимые переменные (предикторы, факторы, регрессоры, экзогенные переменные...) – переменные, значения которых известны.
 - Зависимые переменные (эндогенные переменные, ...) – переменные, значения которых надо определить.

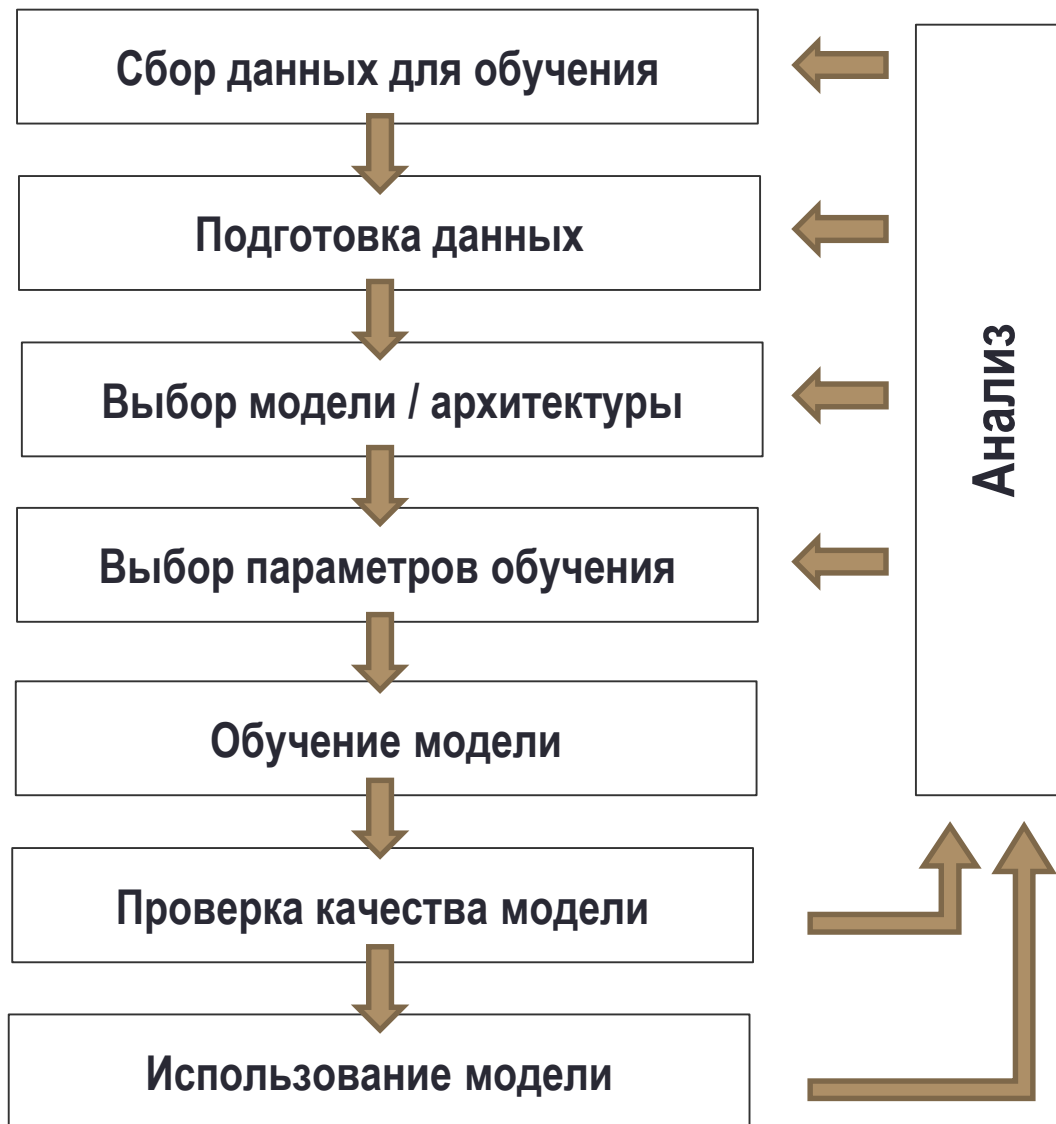
Зависимая переменная	Количественная	Категориальная
Обучение с учителем	Регрессия	Классификация
Обучение без учителя	Уменьшение размерности	Кластеризация

Задачи машинного обучения

- Цель: определить форму функциональной зависимости



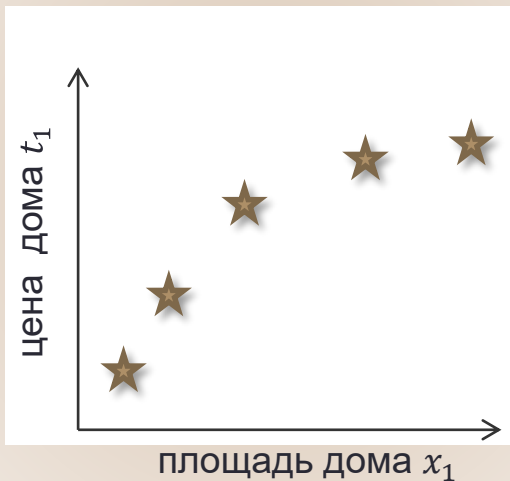
Жизненный цикл модели машинного обучения



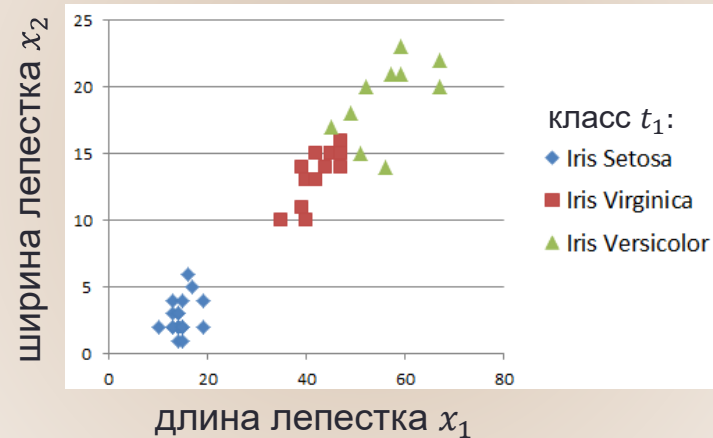
Обучающая выборка

- Каждый элемент обучающей выборки включает:
 - набор значений независимых переменных $\mathbf{x}^{(i)} = x_1^{(i)}, \dots, x_n^{(i)}$
 - соответствующий набор значений зависимых переменных $\mathbf{t}^{(i)} = t_1^{(i)}, \dots, t_m^{(i)}$

Задача регрессии



Задача классификации



Требования к обучающей выборке



● ■ - входят в выборку, ○ □ - не входят в выборку