



MACHINE LEARNING AND AI

ЛЕКЦИЯ 1



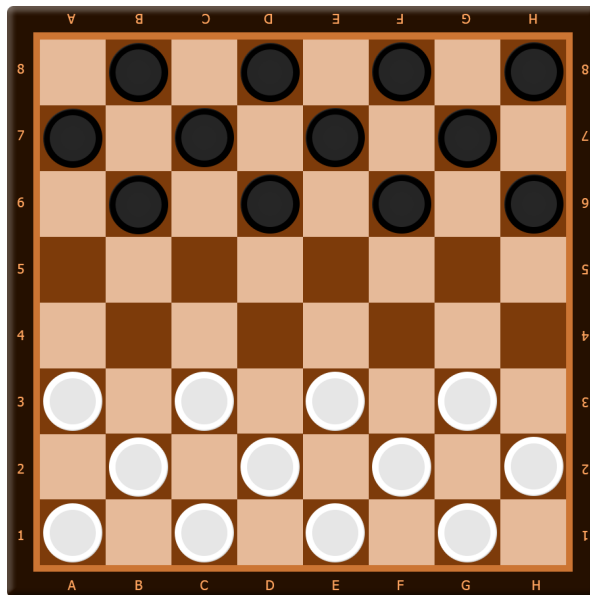
ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Искусственный интеллект – это научное направление

- 1) моделирование процессов познания и мышления
- 2) разработка различных устройств, механизмов, программ, которые по тем или иным критериям могут быть названы «интеллектуальными»
- 3) представления о познании, разуме и человеке, делающих возможным саму постановку вопроса о моделировании интеллекта

МАШИННОЕ ОБУЧЕНИЕ (ML)

- Артур Самюэль (1959). Машинное обучение – область знаний, дающая возможность компьютеру обучаться без строгой программы

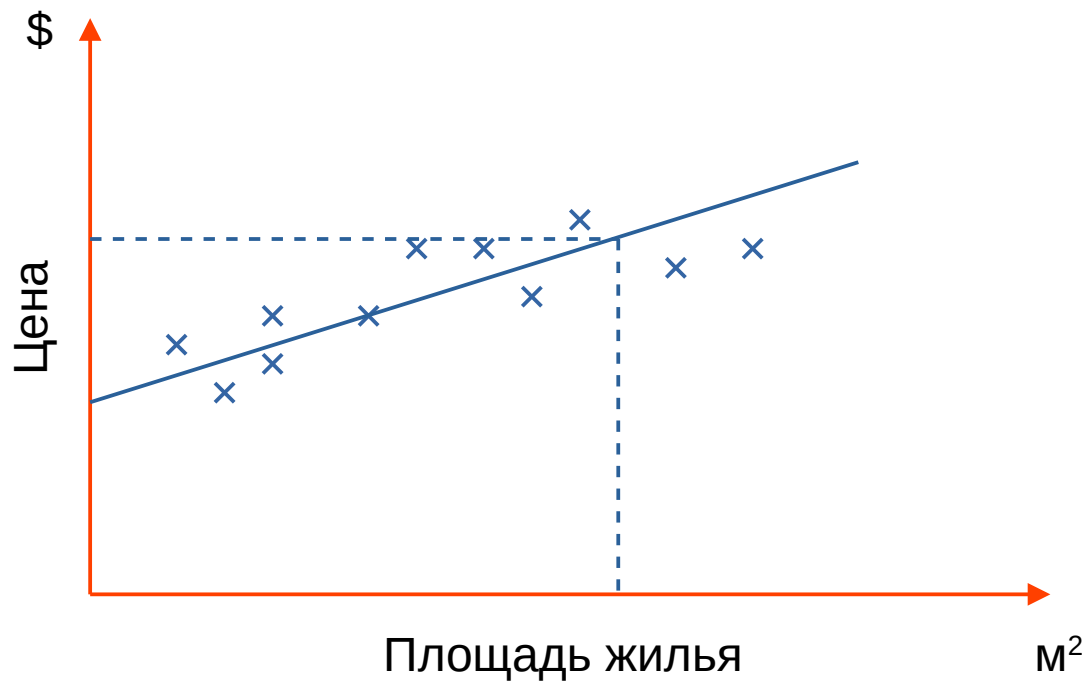




МАШИННОЕ ОБУЧЕНИЕ (ML)

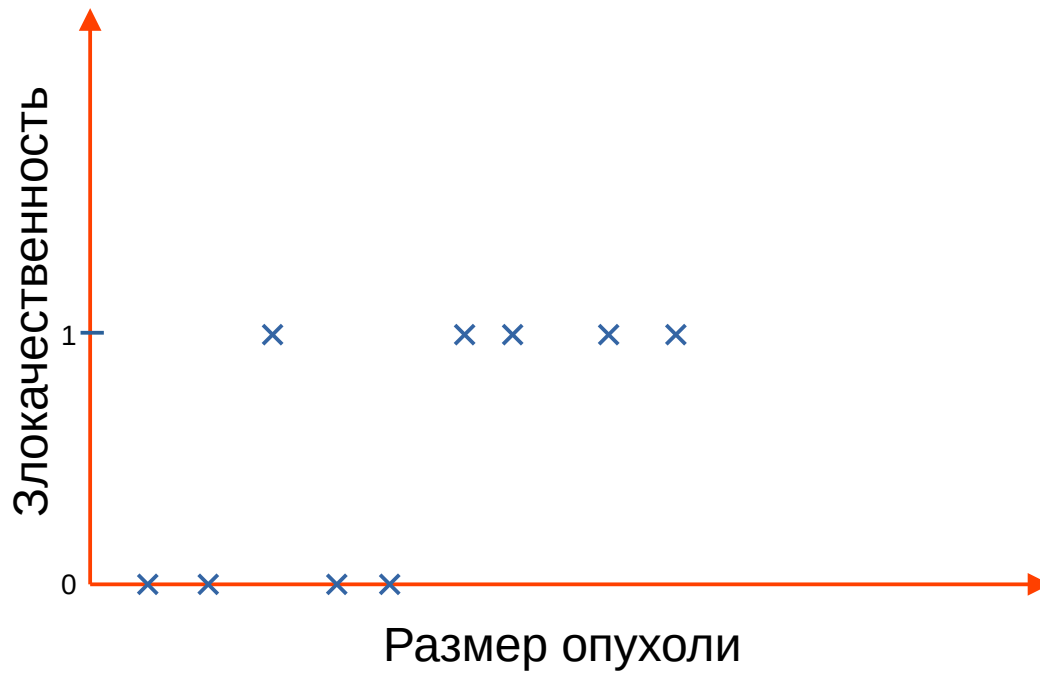
- Артур Самюэль (1959). Машинное обучение – область знаний, дающая возможность компьютеру обучаться без строгой программы
- Том Митчелл (1998). Хорошо поставленная задача обучения – программа, которую можно назвать обучающейся из опыта по отношению к какой-то задаче с некой оценкой эффективности, если эффективность для данной задачи улучшается при наличии опыта

ПРИМЕР ML-ЗАДАЧИ



Задача регрессии

ПРИМЕР ML-ЗАДАЧИ



Задача классификации



КЛАССИФИКАЦИЯ ML-ЗАДАЧ

- Обучение с учителем (Supervised Learning)
 1. Регрессия
 2. Классификация
 3. Анализ временных рядов¹
- Обучение без учителя (Unsupervised Learning)
 1. Кластеризация
 2. Обучение с подкреплением



ПРИМЕРЫ ИЗ ЖИЗНИ

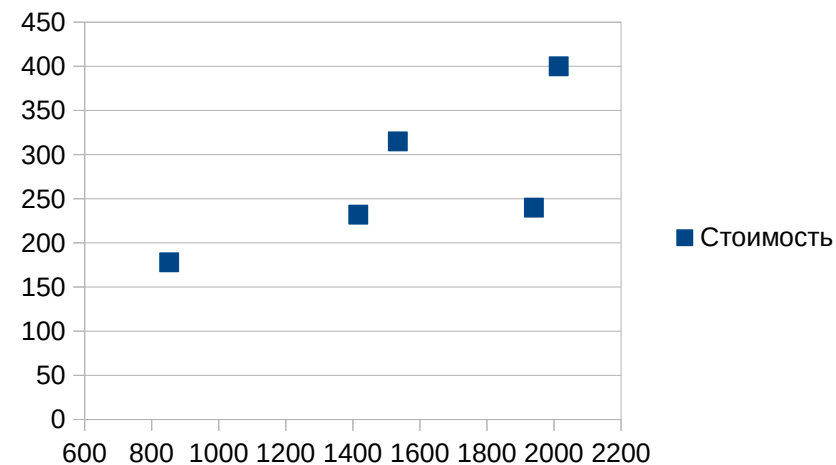
- Банковский скоринг
- Распознавание образов
- Управление машинами и механизмами
- Экспертные системы
- Выработка стратегий
- Приборостроение
- Анализ естественного языка

ПОЛЕЗНЫЕ ССЫЛКИ

- <https://www.youtube.com/playlist?list=PLA89DCFA6ADACE599>
- <https://www.coursera.org/learn/machine-learning>
- <http://www.machinelearning.ru/>
- <https://www.kaggle.com/>
- <https://cs.stanford.edu/people/karpathy/>

ЛИНЕЙНАЯ РЕГРЕССИЯ

Площадь	Число спален	Стоимость
2014	3	400
1416	2	232
1534	3	315
852	2	178
1940	4	240
...





ОБОЗНАЧЕНИЯ

m – число обучающих примеров

X – входные переменные/признаки

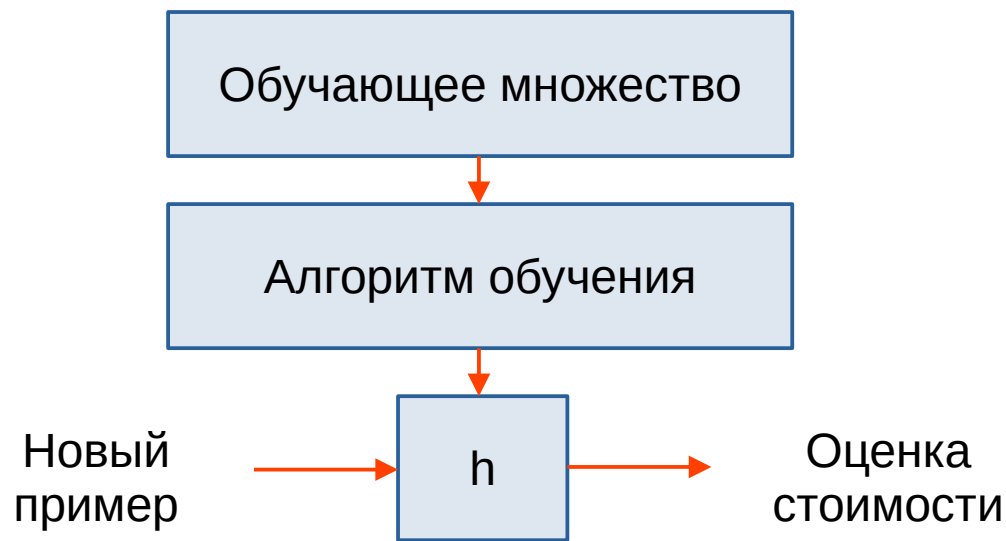
Y – выходная переменная/таргет

(X, Y) – обучающий пример

i -й обучающий пример = $(X^{(i)}, Y^{(i)})$

n – число признаков

СХЕМА РЕШЕНИЯ



$$h(X) = a_1x_1 + a_2x_2 + \dots + a_nx_n$$

$A = (a_1; a_2; \dots ; a_n)$ – параметры

ПОДБОР ПАРАМЕТРОВ

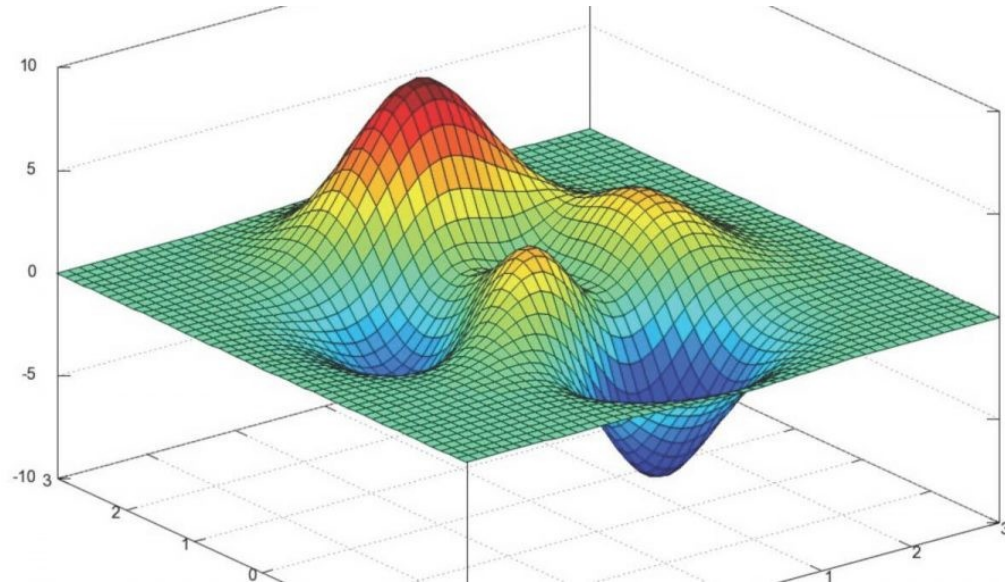
$$\min_A \frac{1}{2} \sum_{i=1}^m (h_A(X) - Y)^2$$

$$J(A) = \frac{1}{2} \sum_{i=1}^m (h_A(X) - Y)^2$$

минимизируем $J(A)$

ГРАДИЕНТНЫЙ СПУСК

- Начинаем с первого приближения A , например $A = [0; 0; \dots; 0]$
- Продолжаем изменять A , уменьшая $J(A)$



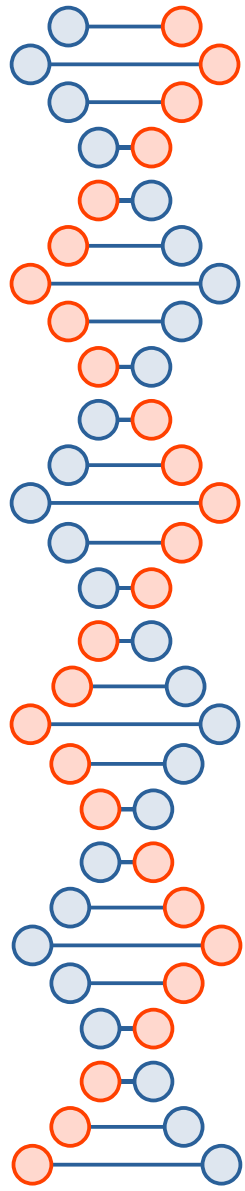
ГРАДИЕНТНЫЙ СПУСК

- Начинаем с первого приближения A , например $A = [0; 0; \dots; 0]$
- Продолжаем изменять A , уменьшая $J(A)$

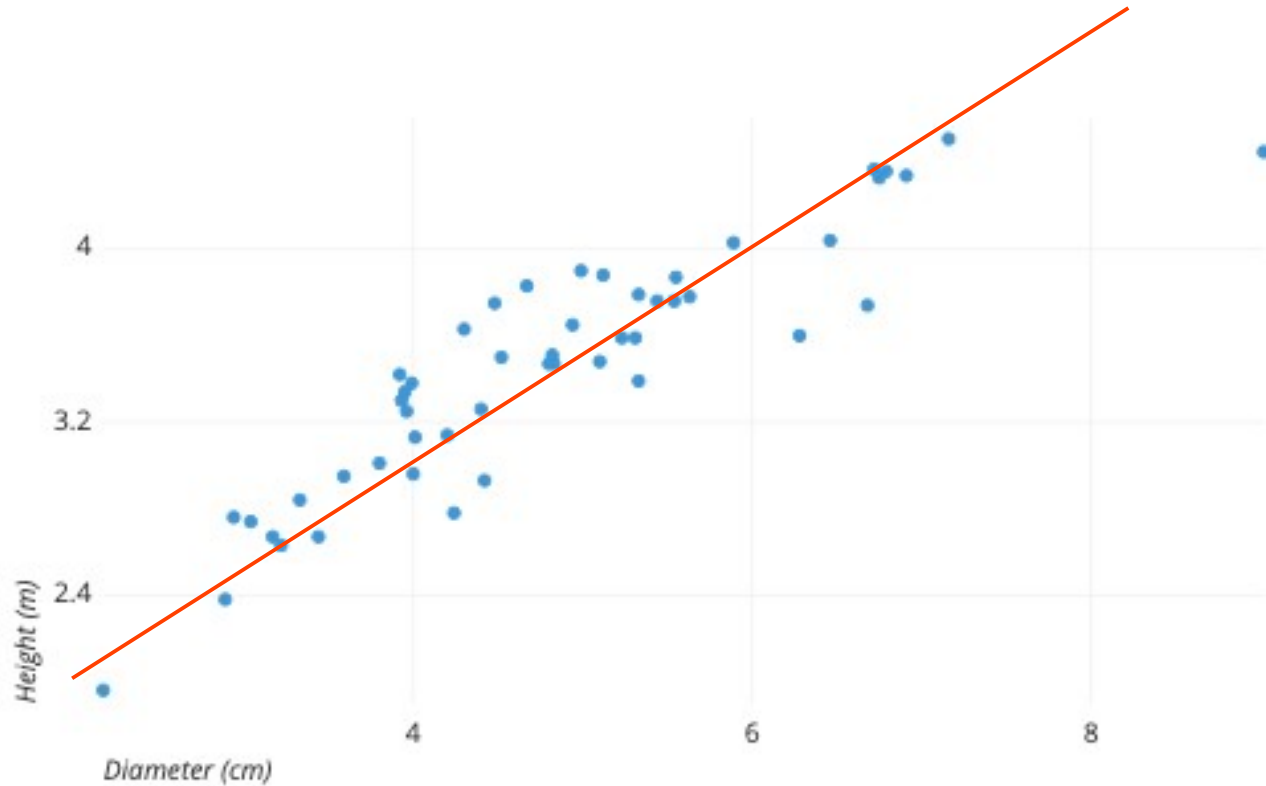
$$a_i := a_i - \alpha \frac{\partial}{\partial a_i} J(A)$$

$$a_i := a_i - \alpha (h_A(X) - Y) x_i$$

$$a_i := a_i - \alpha \sum_{j=1}^m (h_A(X^{(j)}) - Y^{(j)}) x_i^{(j)}$$



ВЕРОЯТНОСТНАЯ ИНТЕРПРЕТАЦИЯ





ВЕРОЯТНОСТНАЯ ИНТЕРПРЕТАЦИЯ

$$y^{(i)} = A^T X^{(i)} + \varepsilon^{(i)}$$

$\varepsilon^{(i)}$ – ошибка

$\varepsilon^{(i)} \sim N(0, \sigma^2)$ – independently identically distributed (IID)

$$L(A) = \prod_{i=1}^m P(y^{(i)} \vee X^{(i)}; A) \quad L(A) = \prod_{i=1}^m \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y^{(i)} - A^T X^{(i)})^2}{2\sigma^2}\right)$$

Максимизируем $L(A)$ - метод максимального правдоподобия

Или минимизируем $-l(A) = -\log(L(A))$