



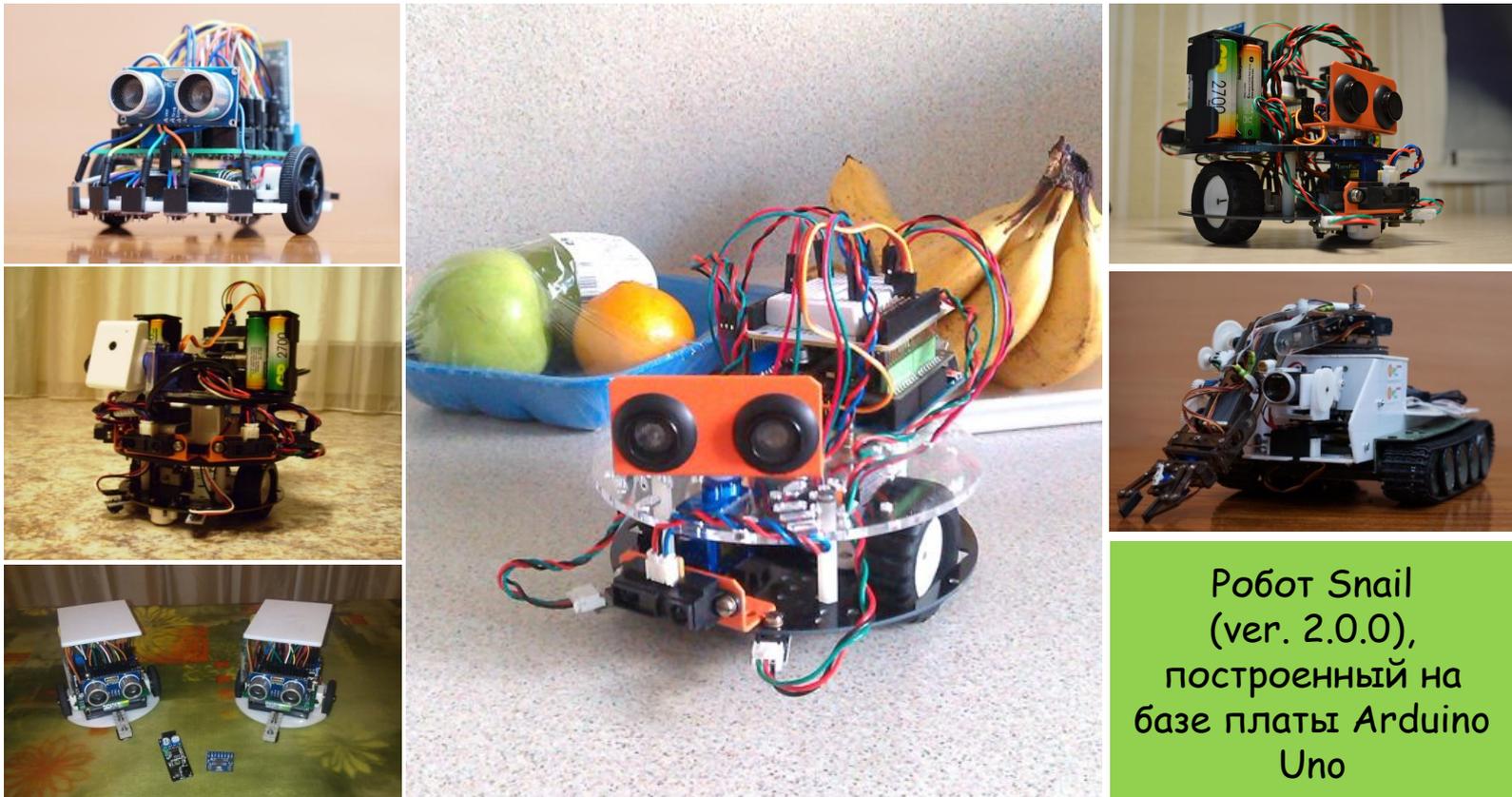
# Прикладная информатика: машинное обучение

Введение и обзор материала курса.  
Линейная регрессия с одной переменной

Владимир Волохов  
физический факультет, лаборатория  
«Цифровые цепи и сигналы»,  
[volokhov.blogspot.com](http://volokhov.blogspot.com)



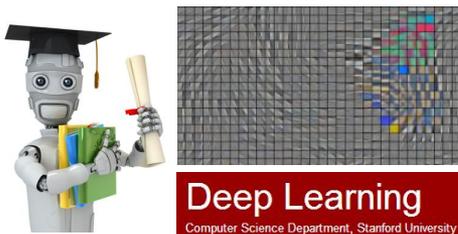
# Машинное обучение



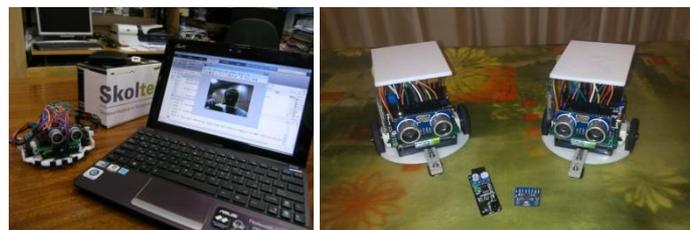
Робот Snail  
(ver. 2.0.0),  
построенный на  
базе платы Arduino  
Uno



# Можно задавать любые вопросы! Мои знания в этой области ограничены ...



Курсы по машинному обучению ([www.coursera.org](http://www.coursera.org), [www.udacity.com](http://www.udacity.com) и т.п.)



Курсовые и дипломные проекты на темы, связанные с машинным обучением



Диссертации и прочее ...

... а также курсы в ФГБОУ ВО ЯрГУ им. П. Г. Демидова («Методы машинного обучения» и «Системы технического зрения»)



# План на сегодня

- ✓ Обзор программы курса
- ✓ Требования
- ✓ Введение в методы машинного обучения
  - ✓ Немного о жизни!
  - ✓ Что такое машинное обучение?
  - ✓ Приложения машинного обучения
  - ✓ Алгоритмы машинного обучения
  - ✓ Линейная регрессия с одной переменной
  - ✓ Метод градиентного спуска
  - ✓ Базы данных



## Цель курса

- ✓ Рассмотрение ряда задач и вопросов из области машинного обучения для:
  - ✓ Создания базиса с целью дальнейшего самообразования
  - ✓ Помощи в дальнейшей практической деятельности как в этой области, так и в смежных областях
  - ✓ Расширения кругозора
- ✓ В общем станет понятнее что это такое и действительно ли вам это нужно



# Структура курса

- ✓ Лекции
  - ✓ Тематика курса охватывает: обучение с учителем, обучение без учителя, регрессия, классификация, кластеризация, сокращение размерности данных, детектирование аномалий, машинное обучение на больших базах данных и глубокое обучение
- ✓ Аттестация по курсу:
  - ✓ Решение практических заданий и тестов по перечисленным темам. Будем использовать язык программирования Python для решения практических заданий!

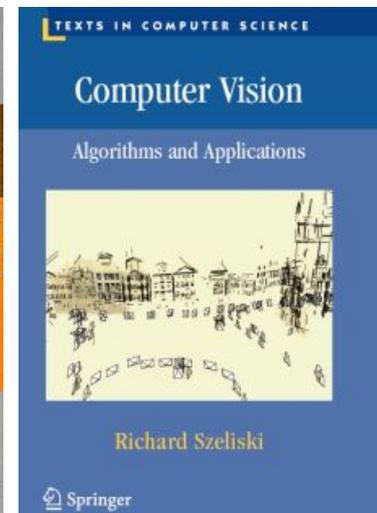
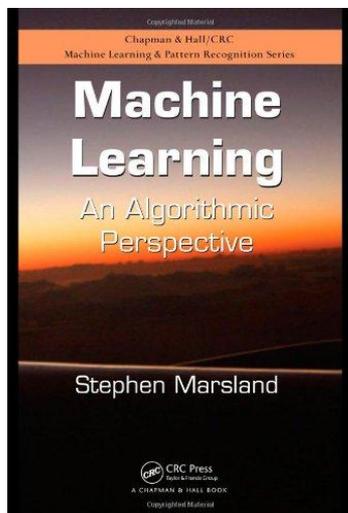


# При решении практических заданий понадобятся

- ✓ Python 3 – высокоуровневый язык программирования общего назначения, ориентированный на повышение производительности разработчика и читаемости кода
- ✓ А также дополнительные библиотеки: NumPy, SciPy, Matplotlib и Scikit-Learn
- ✓ Опционально не будут лишними OpenCV, Pandas и Scikit-Image



# Какую литературу можно найти?



- ✓ Marsland S. Machine Learning: An Algorithmic Perspective, 2009
- ✓ Мерков А.Б. Распознавание образов: введение в методы статистического обучения, 2011
- ✓ Лукьяница А.А., Шишкин А.Г. Цифровая обработка видеоизображений, 2009
- ✓ Шапиро Л., Стокман Дж. Компьютерное зрение, 2006
- ✓ Szeliski R. Computer vision: Algorithm and applications, 2010



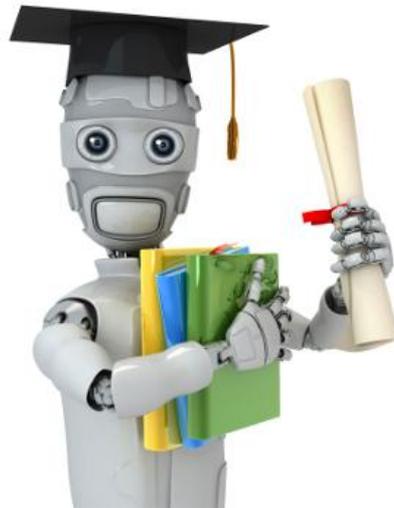
# Какую литературу можно найти?

- ✓ Много информации по теме курса можно найти в сети Интернет
  - ✓ Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных, <http://www.machinelearning.ru>
  - ✓ Большое количество видеолекций по тематике курса можно найти на <http://www.youtube.com> по ключевому слову machine learning
  - ✓ ...



# Немного рекламы!

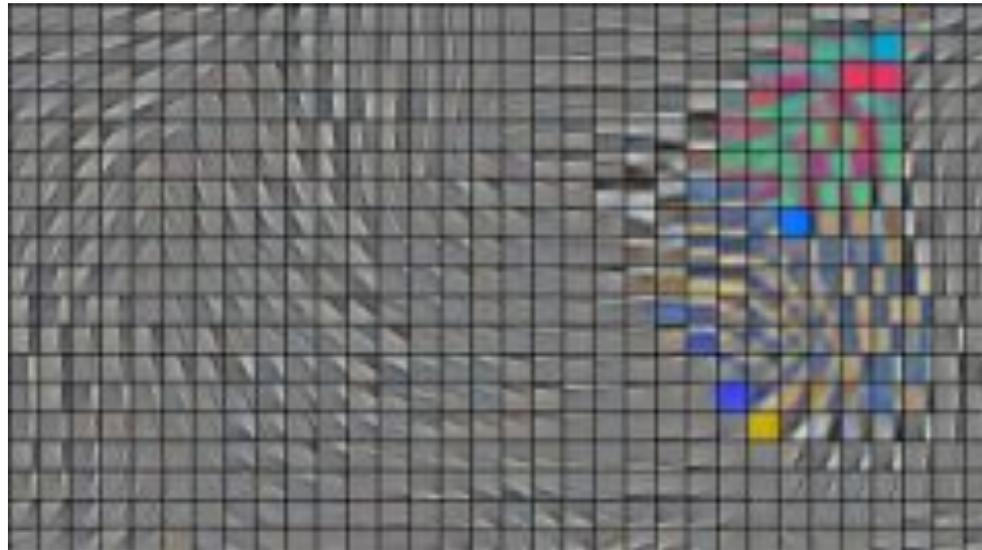
- ✓ Online курс Стэнфордского университета «Машинное обучение»
- ✓ <https://www.coursera.org/learn/machine-learning>





# Немного рекламы!

- ✓ Online курс Торонтского университета «Нейронные сети в машинном обучении»
- ✓ <https://www.coursera.org/learn/neural-networks>





# Немного рекламы!

- ✓ Online курс Высшей школы экономики «Введение в машинное обучение»
- ✓ <https://www.coursera.org/learn/vvedenie-mashinnoe-obuchenie>





# Немного рекламы!

- ✓ Другие курсы по машинному обучению на [www.coursera.org](http://www.coursera.org)
- ✓ Ищем по ключевому слову machine learning. Курсов много, очень много!

 <p>Машинное обучение Стэнфордский университет</p>	 <p>Наука о данных Специализация из 10 курсов · Университет Джона Хопкинса</p>	 <p>Машинное обучение и анализ данных Специализация из 6 курсов · Московский физико-технический институт</p>
 <p>Машинное обучение Специализация из 4 курсов · Вашингтонский университет</p>	 <p>Practical Machine Learning Университет Джона Хопкинса</p>	 <p>Наука о данных для руководителей Специализация из 5 курсов · Университет Джона Хопкинса</p>
 <p>Machine Learning Foundations: A Case Study Approach Вашингтонский университет</p>	 <p>Neural Networks for Machine Learning Торонтский университет</p>	 <p>Обучение на размеченных данных Московский физико-технический институт</p>
 <p>Big Data Специализация из 6 курсов · Калифорнийский университет в Сан-Диего</p>	 <p>Machine Learning: Clustering &amp; Retrieval Вашингтонский университет</p>	 <p>Machine Learning: Classification Вашингтонский университет</p>
 <p>Machine Learning With Big Data Калифорнийский университет в Сан-Диего</p>	 <p>Введение в машинное обучение Высшая школа экономики</p>	 <p>Google Cloud Platform Big Data and Machine Learning Fundamentals Google Cloud</p>



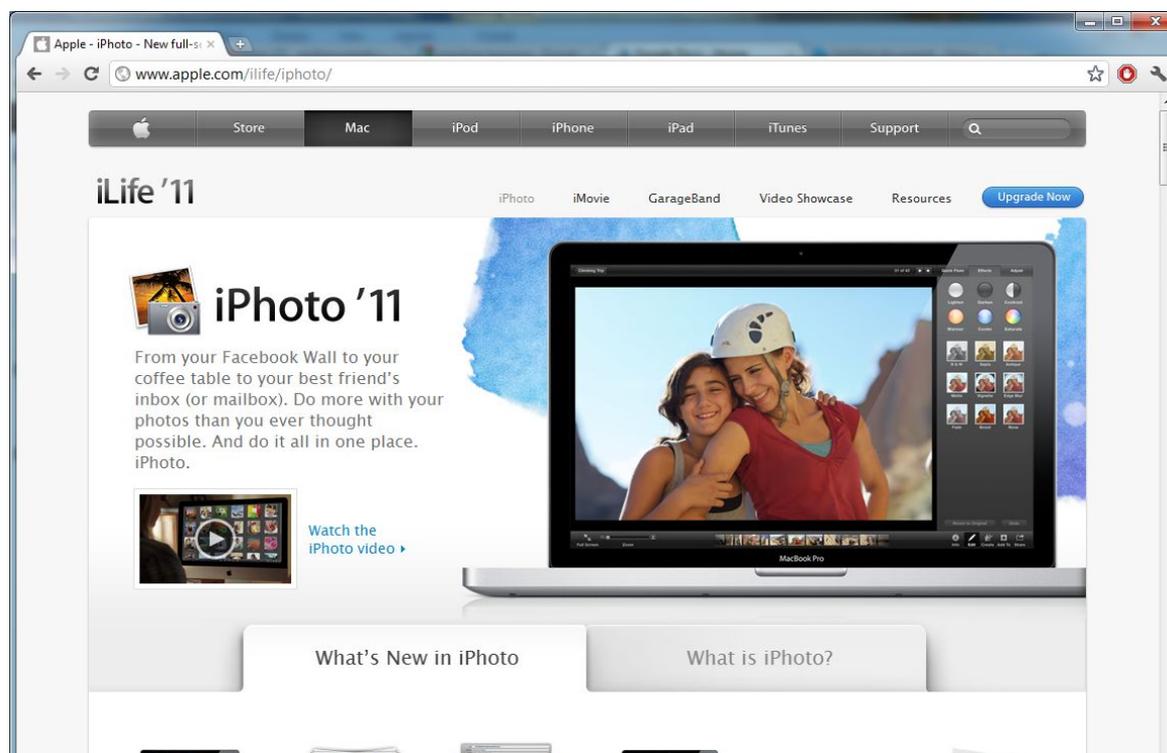
# Немного рекламы!

- ✓ Online курс «Введение в машинное обучение»
  - ✓ <https://www.udacity.com/course/intro-to-machine-learning--ud120>
- ✓ Online курс Технологического института Джорджии «Машинное обучение»
  - ✓ <https://www.udacity.com/course/machine-learning--ud262>
- ✓ Online курс Стэнфордского университета «Введение в искусственный интеллект»
  - ✓ <https://www.udacity.com/course/intro-to-artificial-intelligence--cs271>



# Немного о жизни!

✓ Распознавание лиц: Apple iPhoto software



<http://www.apple.com/ilife/iphoto/>



# Немного о жизни!

✓ Спаму - нет!





# Машинное обучение

- ✓ Выросло из работ по искусственному интеллекту и является одной из перспективных веток его развития
- ✓ Предоставляет новую способность для компьютеров
- ✓ Примеры:
  - ✓ Интеллектуальный анализ баз данных
    - ✓ Большие базы данных с ростом автоматизации/web (рекламные клики, медицинские записи, последовательности ДНК и т.п.)
  - ✓ Приложения, которые не могут быть запрограммированы вручную (автономный вертолет, распознавание рукописного текста и т.п.)
  - ✓ Самозаказывающие программы (Amazon, Netflix и т.п.)
  - ✓ Понимание обучения человека и человеческого мозга



# АВТОНОМНЫЙ ВЕРТОЛЕТ

✓ Создан сотрудниками лаборатории «Искусственного интеллекта» Стэнфордского университета



<http://heli.stanford.edu/>



# Автономное вождение. Проект ALVINN

✓ Снимки сделаны 23 ноября 1992 года!



Autonomous Land Vehicle In a Neural Network (ALVINN) является системой, которая обучается управлять автомобилем путем наблюдения за вождением реального человека. Архитектура ALVINN основана на нейронной сети с одним скрытым слоем, обучаемой с использованием метода обратного распространения ошибки



# Stanley, Junior и Google-мобиль



ALVINN



Stanley



Junior



Google-мобиль



# Суперкомпьютер IBM Watson

✓ В 2011 г. компьютер одержал победу над двумя людьми в телепередаче Jeopardy («Своя игра»), получив один миллион долларов!



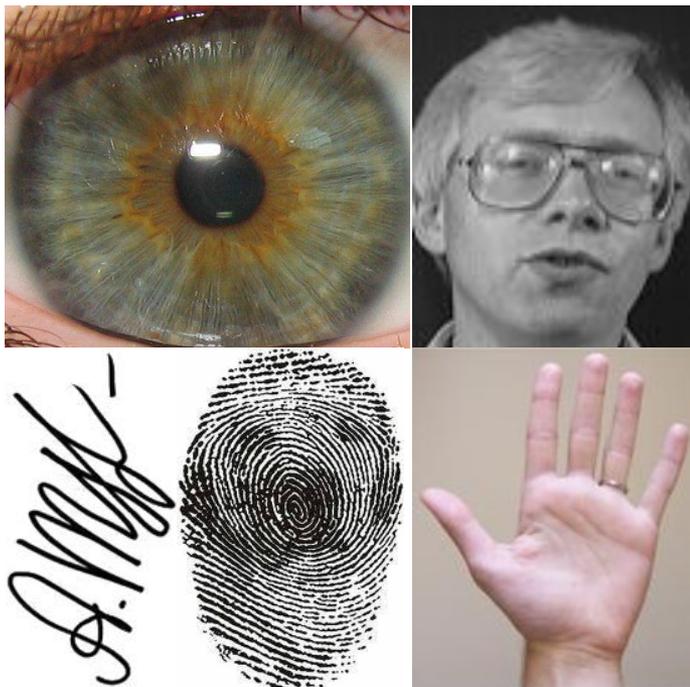
Кен Дженнингс:  
«Ватсон», «Своя игра»  
и я, устаревший  
всезнайка

[http://www.ted.com/talks/ken\\_jennings\\_watson\\_jeopardy\\_and\\_me\\_the\\_obsolete\\_know\\_it\\_all?language=ru](http://www.ted.com/talks/ken_jennings_watson_jeopardy_and_me_the_obsolete_know_it_all?language=ru)



# Компьютерное зрение

- ✓ Распознавание лиц, подписи, отпечатков пальцев, радужной оболочки глаза



Иванов  
Иван Иванович



# Компьютерное зрение

## ✓ Зрение для роботов



AiVo (японская собака-робот, изобретение компании Sony)

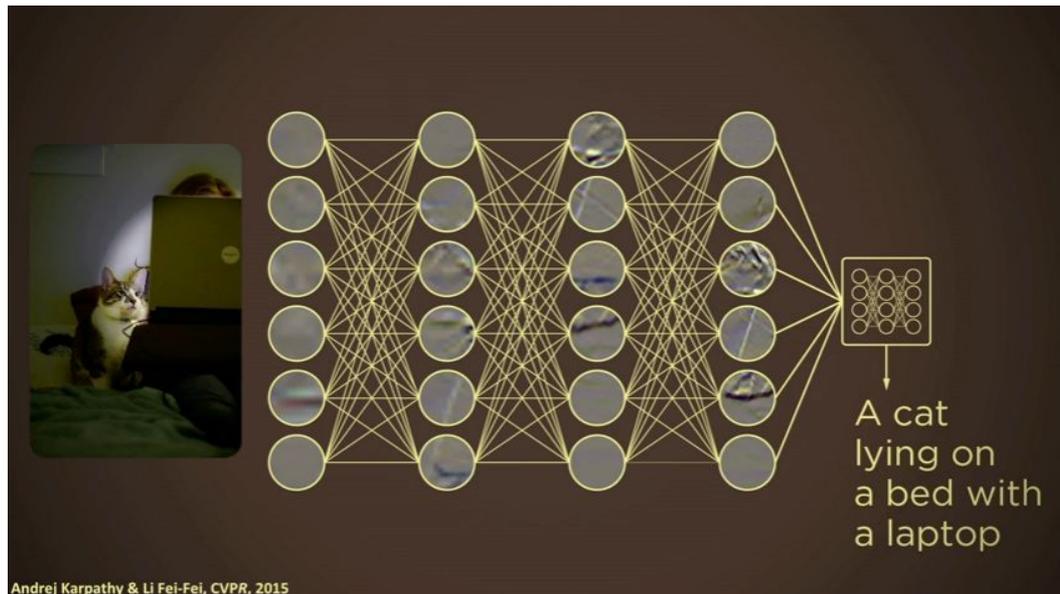


Teddy (японский медведь-робот, изобретение компании Fujitsu)



# Компьютерное зрение

✓ Ли Ф.-Ф. Как мы учим компьютер понимать изображения?



[https://www.ted.com/talks/fei\\_fei\\_li\\_how\\_we\\_re\\_teaching\\_computers\\_to\\_understand\\_pictures?language=en#t-5181](https://www.ted.com/talks/fei_fei_li_how_we_re_teaching_computers_to_understand_pictures?language=en#t-5181)



# Компьютерное зрение

- ✓ Ли Ф.-Ф. Как мы учим компьютер понимать изображения?

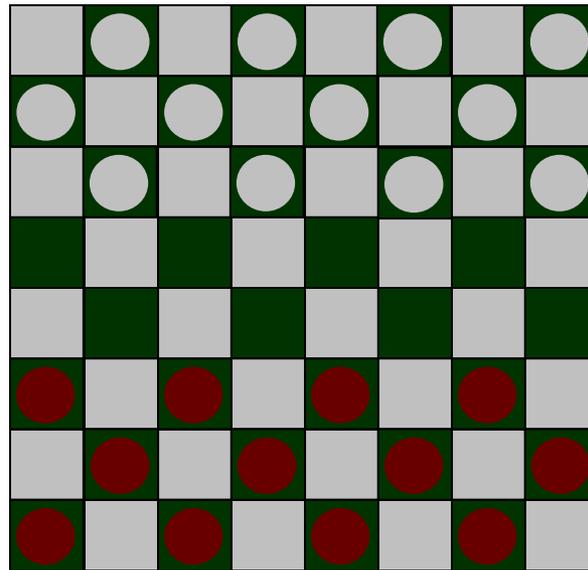


[https://www.ted.com/talks/fei\\_fei\\_li\\_how\\_we\\_re\\_teaching\\_computers\\_to\\_understand\\_pictures?language=en#t-5181](https://www.ted.com/talks/fei_fei_li_how_we_re_teaching_computers_to_understand_pictures?language=en#t-5181)



# Что такое машинное обучение?

✓ А. Самуэль (Arthur Samuel), 1959 г.: **Машинное обучение** - это область исследования, которая предоставляет компьютеру, не будучи явно запрограммированным, возможность к обучению





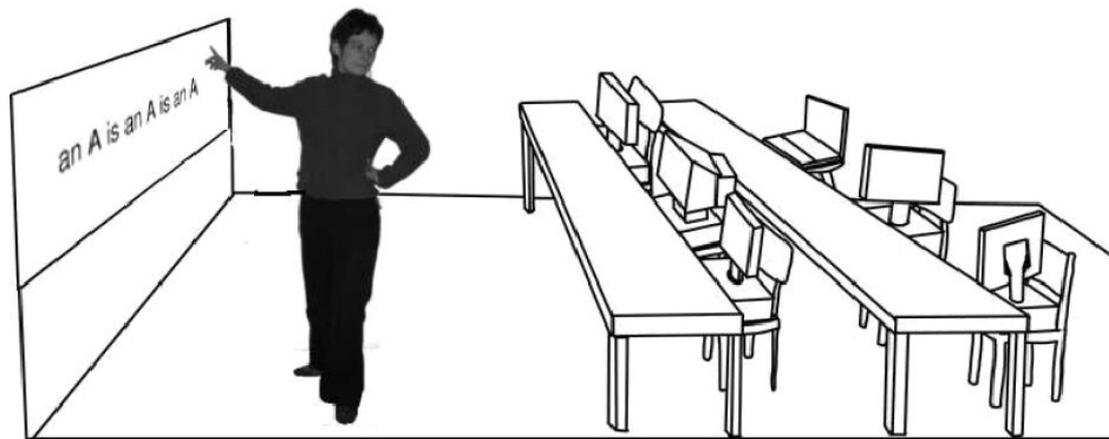
# Что такое машинное обучение?

- ✓ **А. Самуэль (Arthur Samuel), 1959 г.: Машинное обучение** - это область исследования, которая предоставляет компьютеру, не будучи явно запрограммированному, возможность к обучению
- ✓ **Т. Митчелл (Tom Mitchell), 1998 г.:** конкретно поставленная задача обучения. Говорят, что компьютерная программа способна к обучению из опыта  $E$  по отношению к некоторой задаче  $T$  и критерию качества работы  $P$ , если ее производительность на  $T$ , как мера  $P$ , улучшается с опытом  $E$



# Что такое машинное обучение?

- ✓ Обучение = приспособление
  - ✓ Обучаясь, мы становимся более приспособленными к окружающей среде
- ✓ Обучение ≠ «заучивание наизусть»
  - ✓ Заучить наизусть - для машины не проблема
  - ✓ Мы хотим научить машину делать выводы!





# Алгоритмы машинного обучения

- ✓ **Обучение с учителем (Supervised Learning).** Алгоритм машинного обучения обобщает на основе размеченного множества тренировочных примеров
- ✓ **Обучение без учителя (Unsupervised Learning).** Размеченное тренировочное множество отсутствует. Алгоритм машинного обучения пытается самостоятельно определить сходство между входами. Поэтому входы, которые имеют что-то общее объединяются в общую категорию



# Обучение с учителем

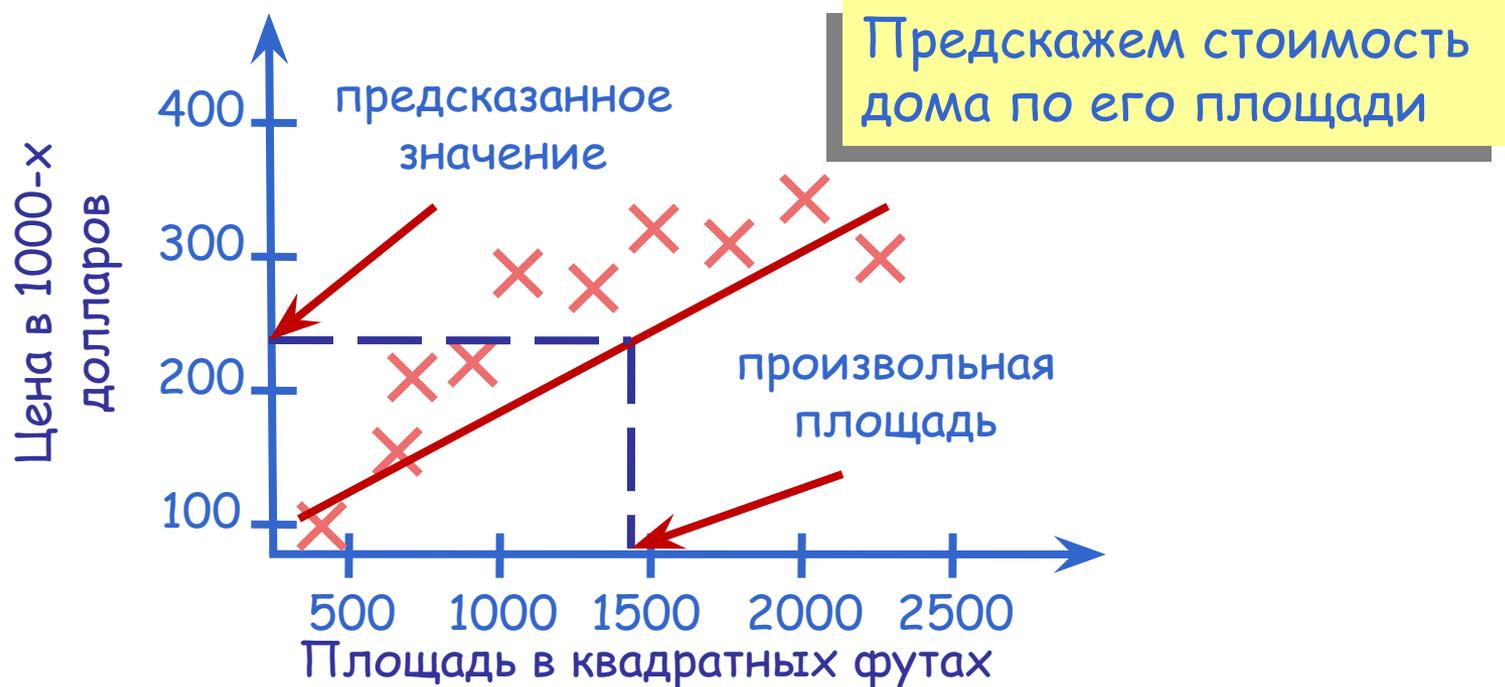
- ✓ Регрессия (предсказание непрерывной выходной величины, например, цены на недвижимость)





# Обучение с учителем

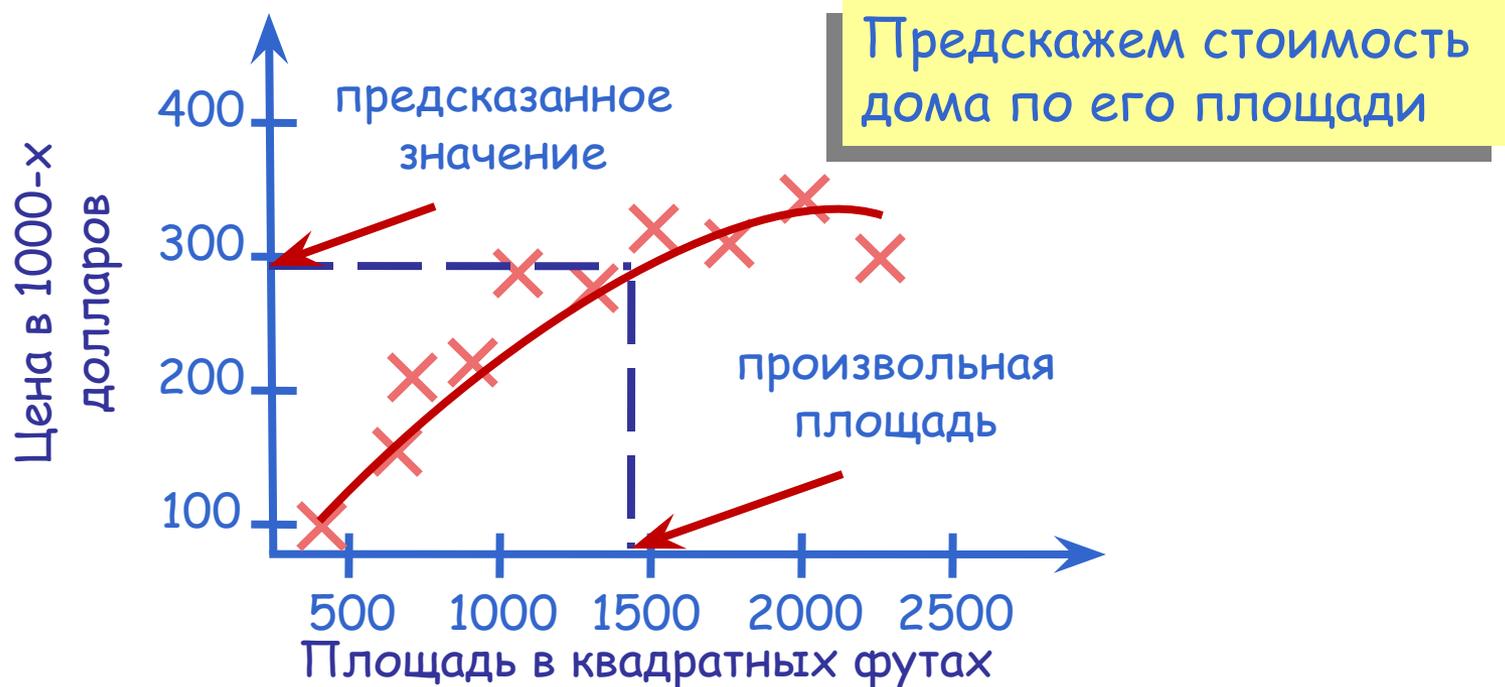
- ✓ Регрессия (предсказание непрерывной выходной величины, например, цены на недвижимость)





# Обучение с учителем

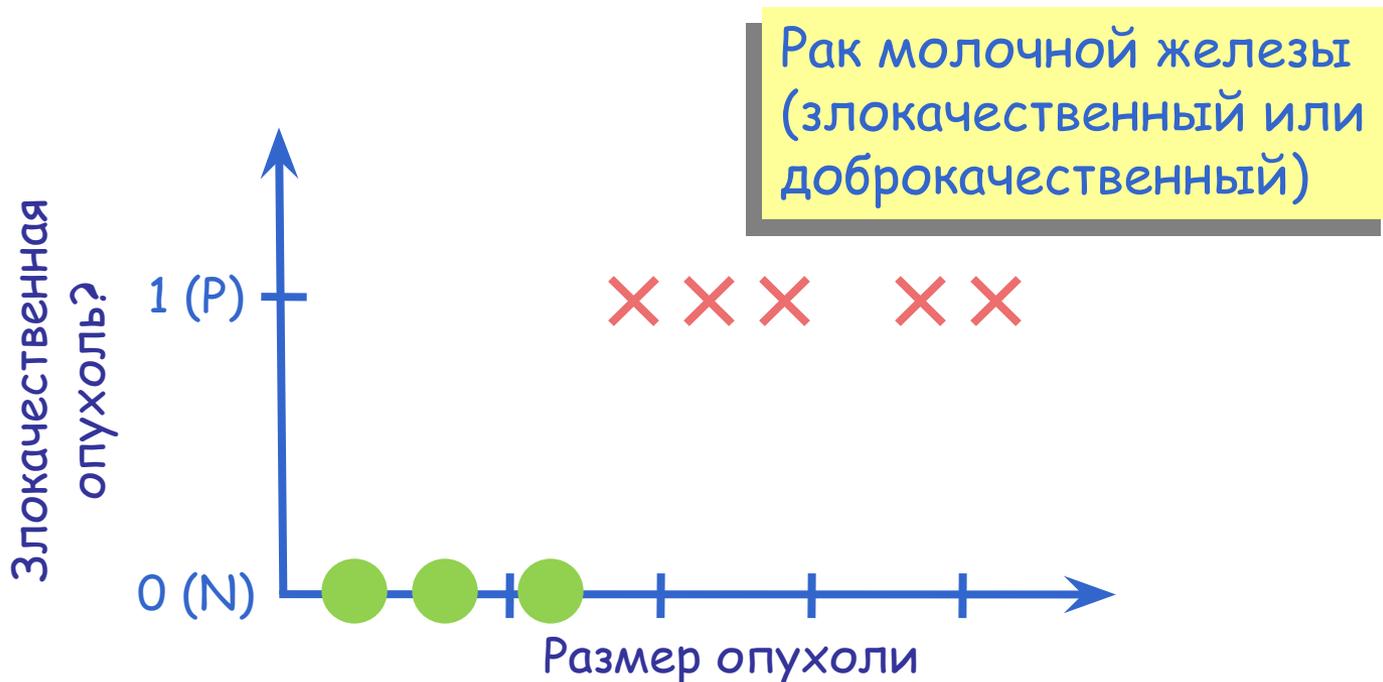
- ✓ Регрессия (предсказание непрерывной выходной величины, например, цены на недвижимость)





# Обучение с учителем

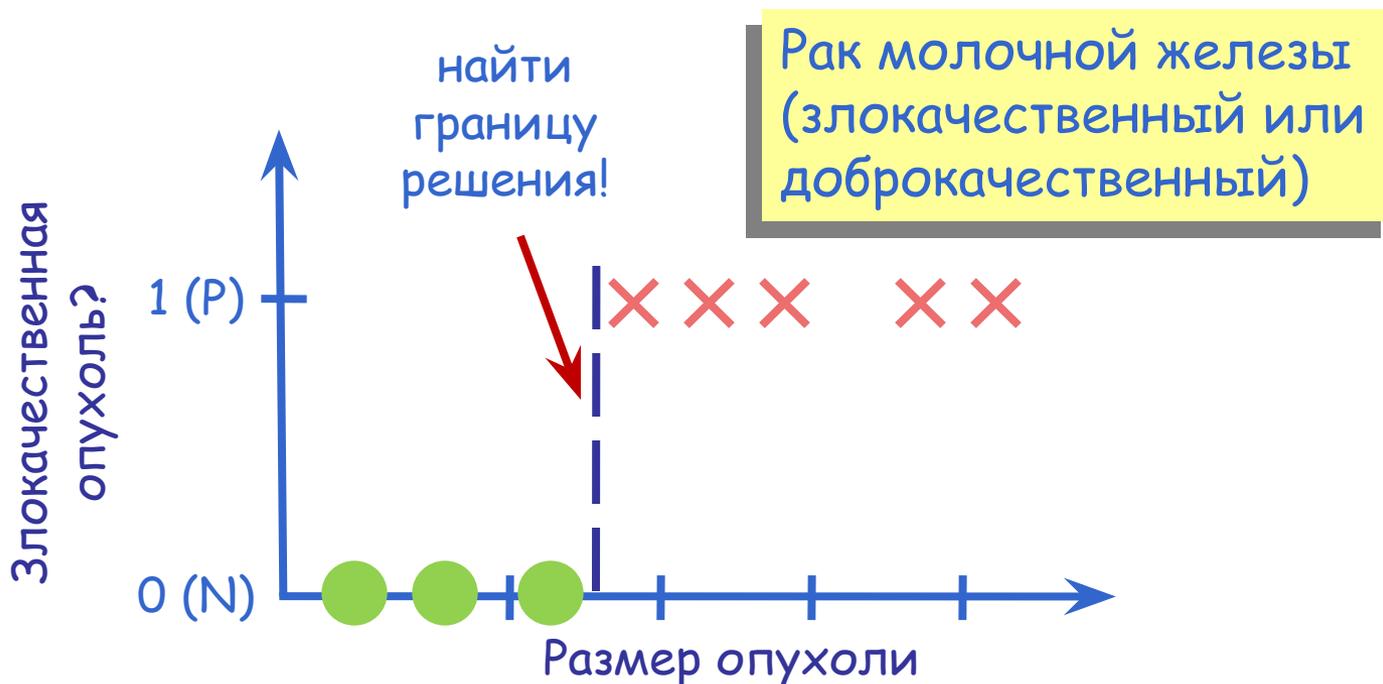
- ✓ Классификация (предсказание дискретной выходной величины, например, 0 или 1)





# Обучение с учителем

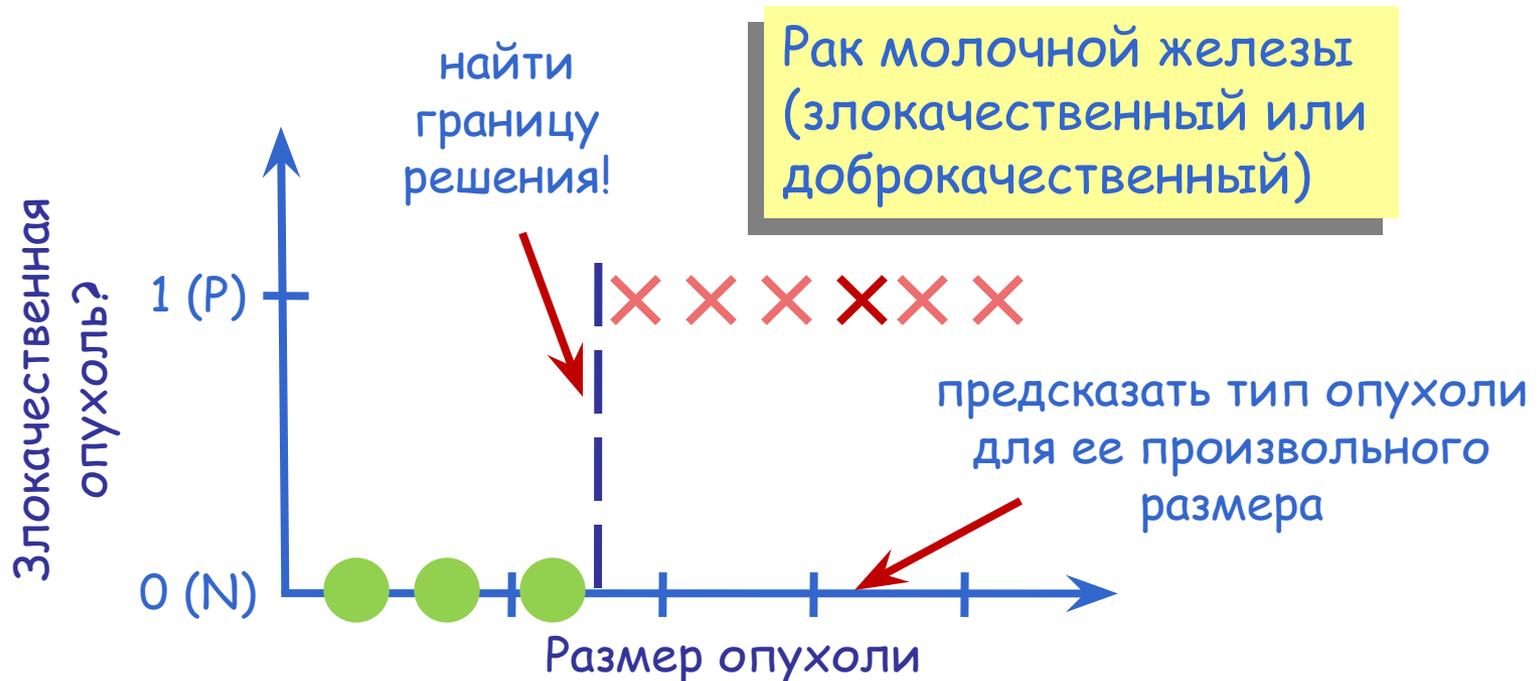
- ✓ Классификация (предсказание дискретной выходной величины, например, 0 или 1)





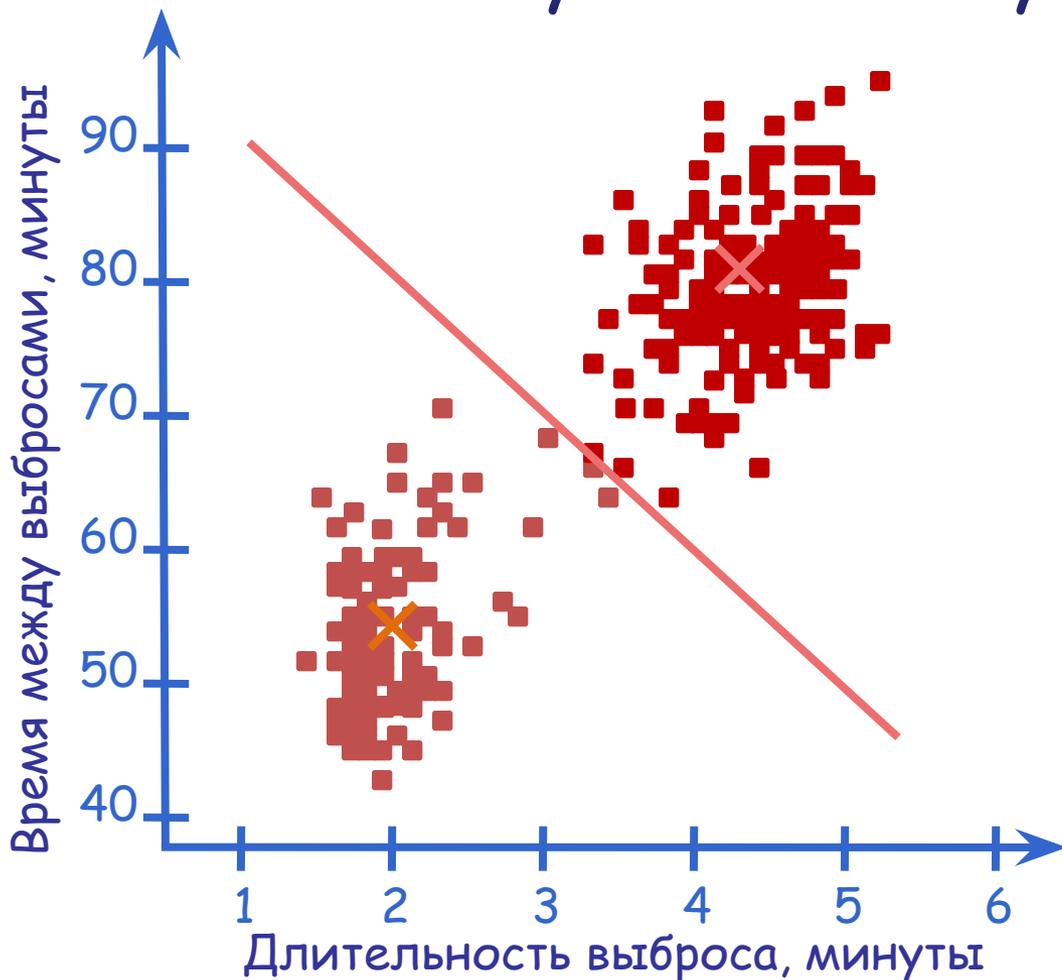
# Обучение с учителем

- ✓ Классификация (предсказание дискретной выходной величины, например, 0 или 1)





## Обучение без учителя

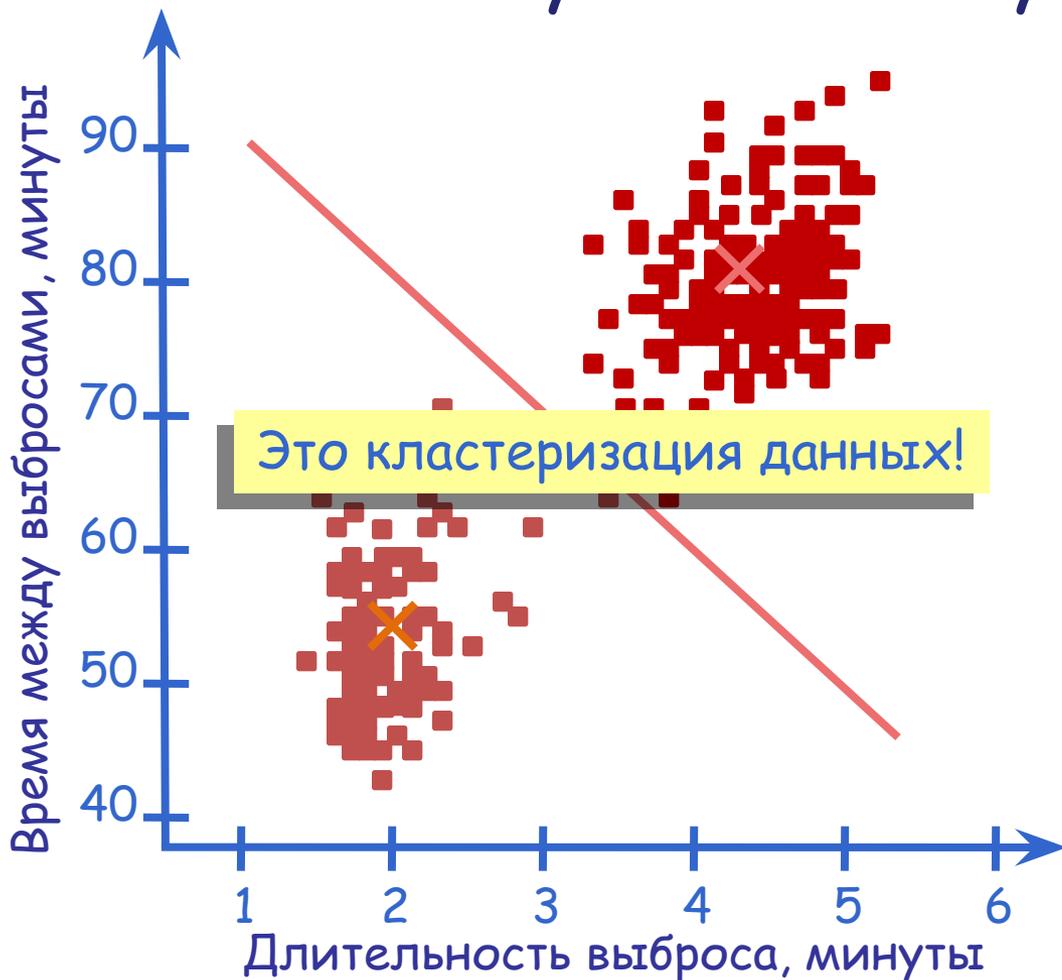


Рассмотрим метод  
K-средних и задачу  
с гейзером





## Обучение без учителя



Рассмотрим метод  
K-средних и задачу  
с гейзером





# Кластеризация в Google News

The screenshot shows the Google News homepage in a browser window. The address bar displays 'news.google.com'. The page features a search bar with 'Search News' and 'Search the Web' buttons. On the left, there is a sidebar with 'Top Stories' and 'All news' categories. The main content area is divided into 'Top Stories' and 'Recent' sections. The 'Top Stories' section includes articles such as 'White House official denies Tea Party-focused ad campaign', 'US Stocks Climb After Recession Called Over, Homebuilders Gain', and 'BP Oil Well, Site of National Catastrophe, Dies at One'. The 'Recent' section includes 'Recession officially ended in June 2009', 'Hurricane Igor lashes Bermuda', and 'Explain what you want from us.' reads front-page editorial. The page also shows a 'Spotlight' section with the article 'Sarkozy rages at EU humiliation'.



# Кластеризация в Google News

The screenshot shows a web browser window displaying a page from The Wall Street Journal's 'The Source' section. The browser's address bar shows the URL: `blogs.wsj.com/source/2010/09/20/bp-kills-macondo-but-its-legacy-lives-on/`. The page header includes 'THE WALL STREET JOURNAL. Digital Network' and navigation links for 'WSJ.com', 'MarketWatch', 'BARRON'S', 'All Things Digital.', 'FINS', 'SmartMoney', and 'More'. A search bar is located in the top right corner.

The main content area features the title 'BP Kills Macondo, But Its Legacy Lives On' and a large image of an offshore oil rig on fire. The article is by James Herron and is dated September 20, 2010. The text describes the Macondo well leak and the impact on BP and the oil industry. A caption below the image reads: 'Fire boat response crews battled the blazing remnants of the off shore oil rig Deepwater Horizon on April 21, 2010.' The image is credited to 'Associated Press'.

On the right side of the page, there is a 'Search The Source' bar, a section titled 'About The Source' with a photo of Lauren Mills, and a 'Most Recent' list of articles. The list includes: 'Who Needs Plaza II Anyway', 'Will Banks Be Forced to Split Retail And Banking Arms?', 'Timing of Ratings Award Intriguing', 'BP Kills Macondo, But Its Legacy Lives On', and 'We Already Need a Sequel to Basel III'.



# Кластеризация в Google News

The screenshot shows a web browser window displaying a CNN article. The address bar shows the URL: `edition.cnn.com/2010/US/09/20/gulf.oil.disaster/`. The page features the CNN logo and navigation menus for various regions and topics. The main headline is "Allen: Well is dead, but much Gulf Coast work remains" by the CNN Wire Staff, dated September 20, 2010. Below the headline is a video player with a "Click to play" button and a play icon. To the right of the video is a "Most Popular" section listing five trending stories, including "White House official denies Tea Party-focused ad campaign" and "Swedish far-right leader: Success due to immigration backlash". At the bottom, a "STORY HIGHLIGHTS" section begins with the text: "(CNN) -- The ruptured Macondo well, a mile under the Gulf of Mexico off the Louisiana coast, has been pronounced dead".



# Кластеризация в Google News

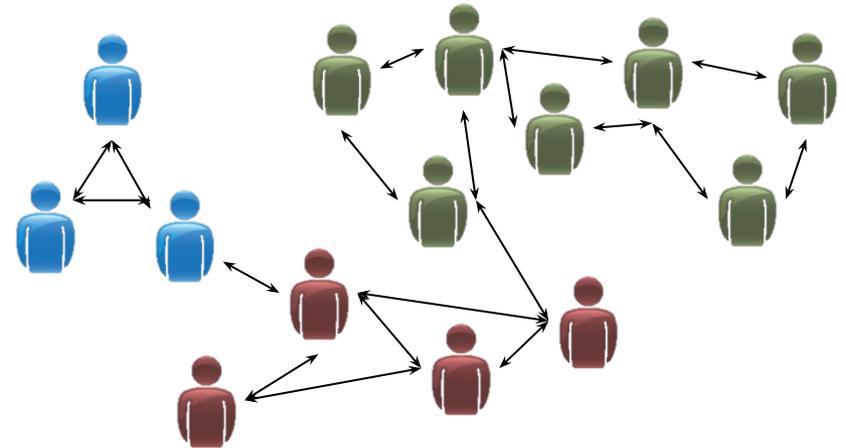
The screenshot shows a web browser window displaying a Guardian article. The URL is [www.guardian.co.uk/environment/2010/sep/20/bp-oil-spill-deepwater-horizon-costs-10bn](http://www.guardian.co.uk/environment/2010/sep/20/bp-oil-spill-deepwater-horizon-costs-10bn). The article title is "BP oil spill cost hits nearly \$10bn". The text states: "BP has set up a \$20bn compensation fund after the Deepwater Horizon disaster, which has so far paid out 19,000 claims totalling more than \$240m". The author is Julia Kollwe, and the article was published on Monday 20 September 2010 08.33 BST. A large photograph shows an offshore oil rig on fire with thick black smoke billowing into the sky, surrounded by support vessels. The article includes social media sharing options (Twitter, Facebook, etc.), a "Comments (10)" section, and a "BP oil spill news on Twitter" sidebar with tweets from BP\_America and guardianeco. At the bottom, there is an "On Business" section with a sub-section "Last 24 hours" featuring a photo of a man and the headline "1. Britain keeps AAA credit rating".



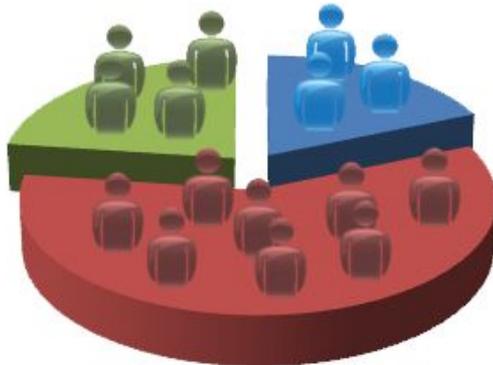
# Еще несколько примеров кластеризации!



Организация компьютерных кластеров



Анализ социальных сетей



Сегментация рынка



Анализ астрономических данных



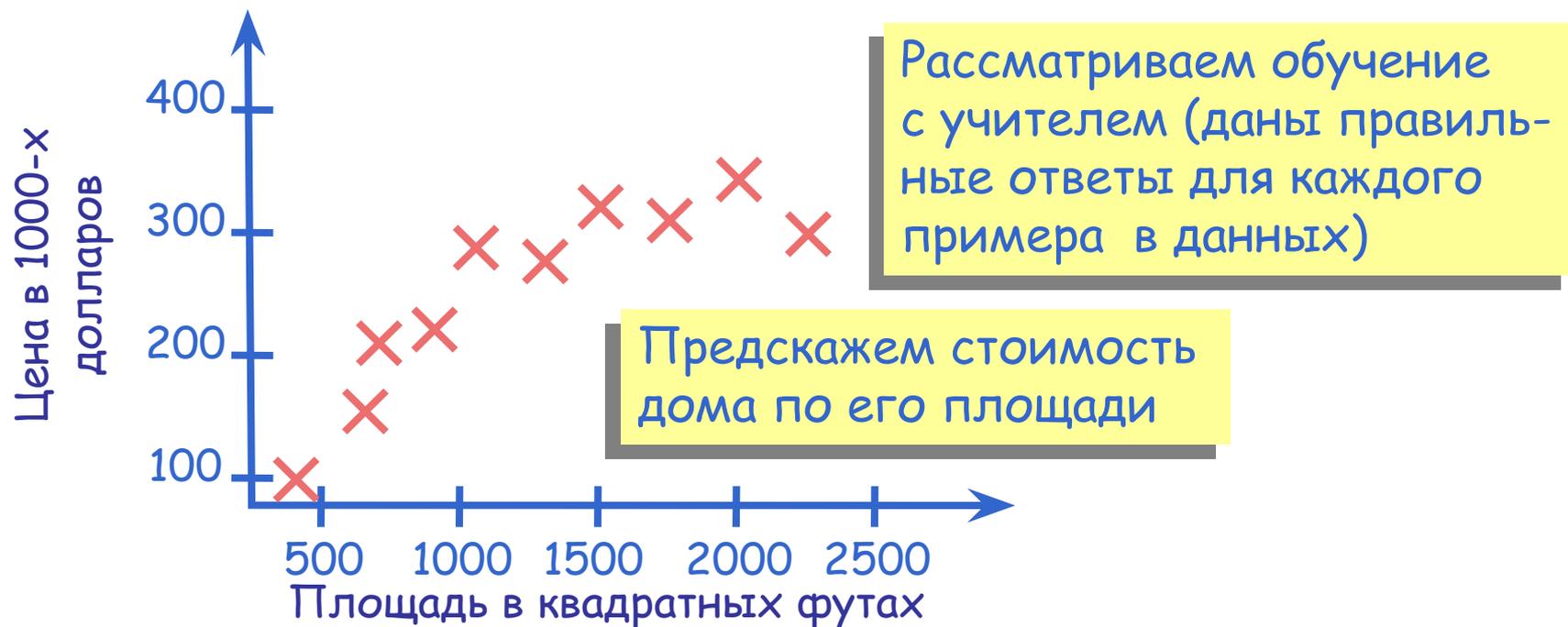
# Алгоритмы машинного обучения

- ✓ **Обучение по алгоритму типа «кнут и пряник» или стимулированное обучение (Reinforcement Learning).** Алгоритм машинного обучения получает подсказку, когда его ответ является ошибочным и не получает, когда ответ является правильным
- ✓ **Эволюционное обучение (Evolutionary Learning).** Биологические организмы адаптируются для того, чтобы улучшить свою выживаемость и шансы иметь потомство в окружающей их среде. Смоделируем подобное на компьютере!



# Линейная регрессия с одной переменной

- ✓ Регрессия (предсказание непрерывной выходной величины, например, цены на недвижимость)





# Линейная регрессия с одной переменной

✓ Тренировочное множество данных (скажем, всего  $m$ )

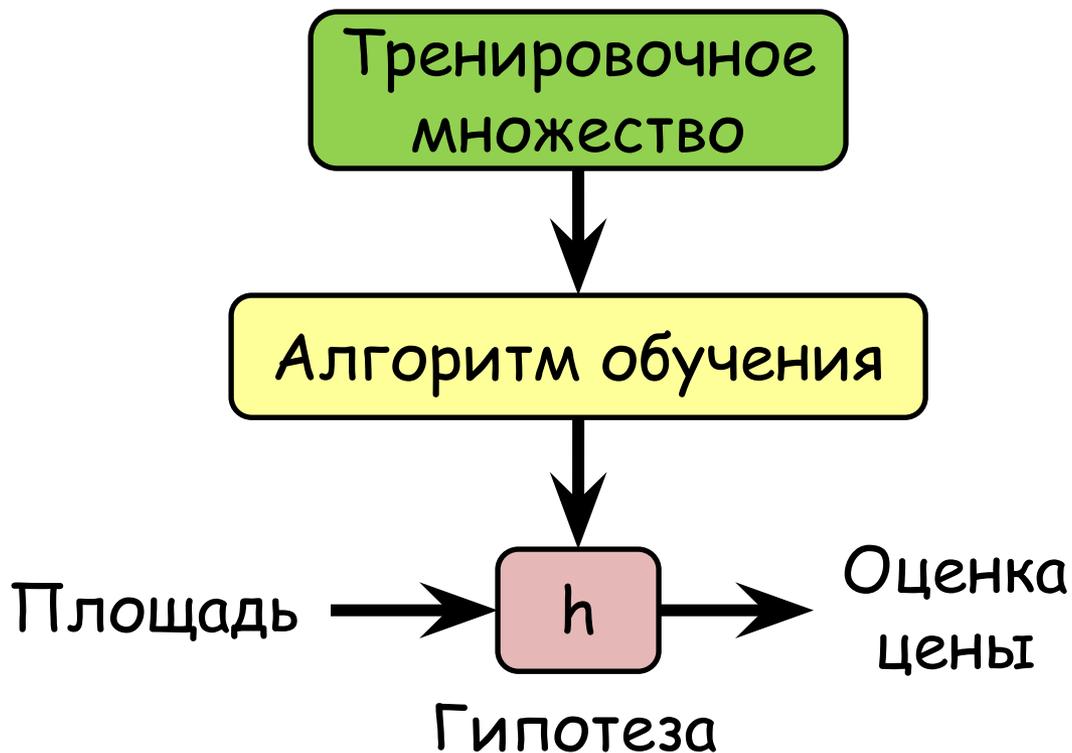
Площадь (фут <sup>2</sup> ) - $x$	Цена в 1000-х (\$) - $y$
2104	460
1416	232
1534	315
852	178
...	...

Обозначения:  $m$  = число тренировочных примеров  
 $x$  = «входная» переменная / свойство  
 $y$  = «выходная» переменная / «метка»  
 $(x^{(i)}, y^{(i)})$  =  $i$ -й тренировочный пример



# Линейная регрессия с одной переменной

## ✓ Последовательность действий



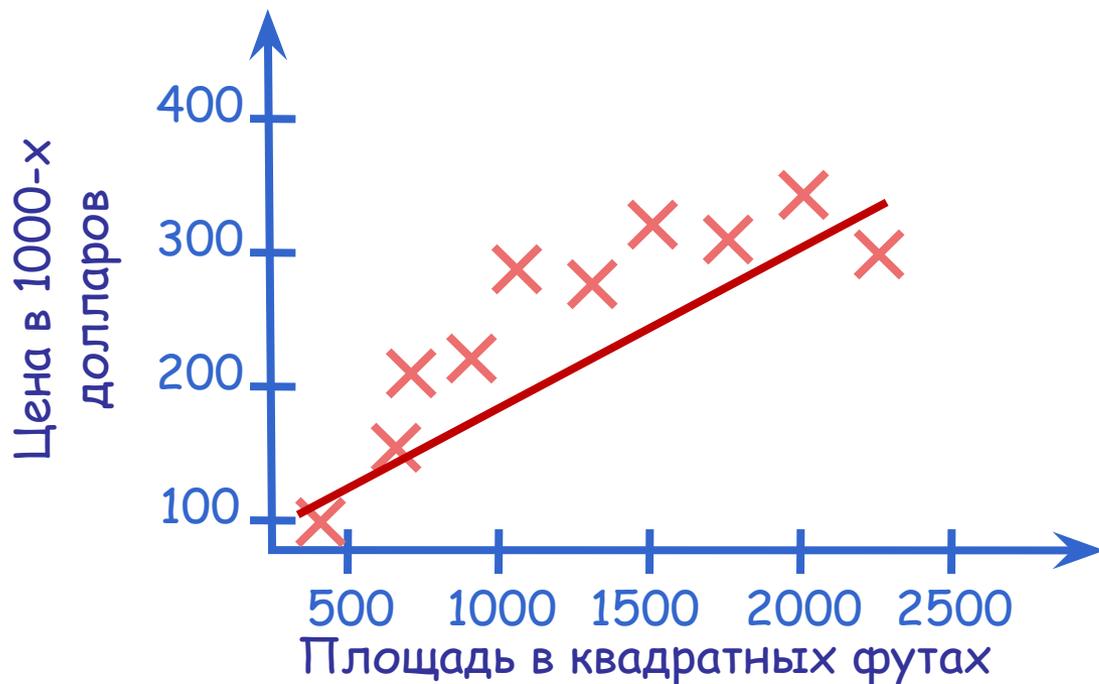
Гипотеза  $h$  выглядит так:  $h_Q(x) = Q_0 + Q_1 x$

Как оценить  $Q_0$  и  $Q_1$ ?



# Линейная регрессия с одной переменной

## ✓ Последовательность действий



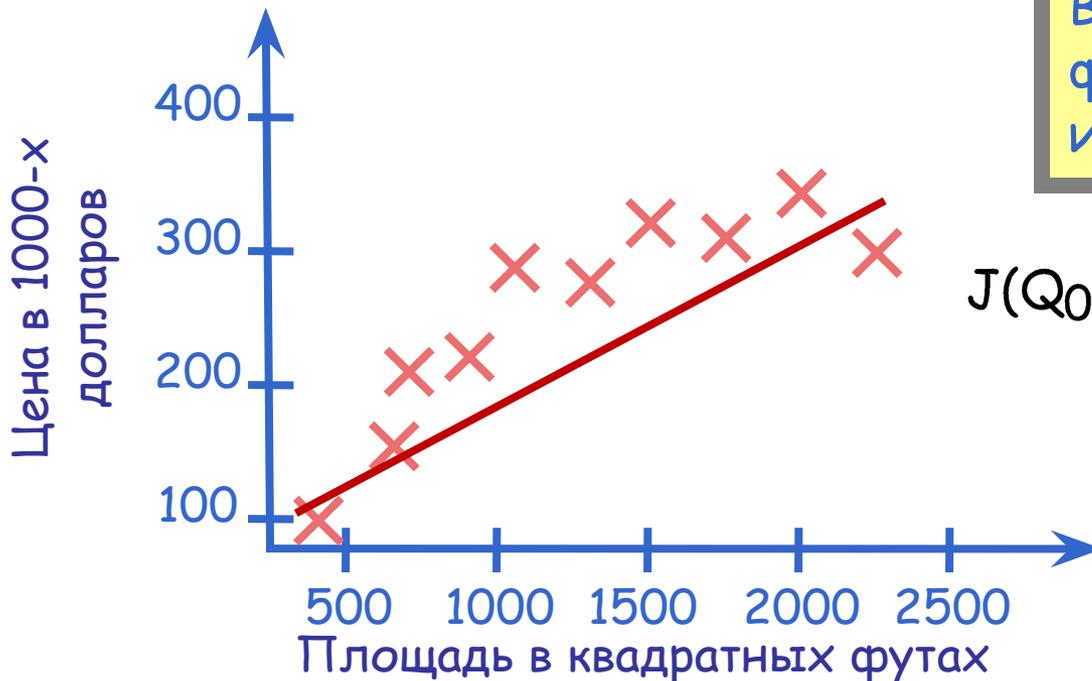
Гипотеза  $h$  выглядит так:  $h_Q(x) = Q_0 + Q_1 x$

Как оценить  $Q_0$  и  $Q_1$  ?



# Стоимостная функция (Cost Function)

Идея! Выбрать  $Q_0$  и  $Q_1$  так, чтобы  $h_Q(x)$  являлась близкой к значениям  $y$  для всех тренировочных примеров



Введем стоимостную функцию  $J(Q_0, Q_1)$  и минимизируем ее:

$$J(Q_0, Q_1) = \frac{1}{2m} \sum_{i=1}^m (h_Q(x^{(i)}) - y^{(i)})^2,$$

$$\min_{Q_0, Q_1} J(Q_0, Q_1)$$



## Как минимизировать стоимостную функцию для линейной регрессии с одной переменной?

- ✓ Использование нормальных уравнений (normal equation)
  - ✓ Рассмотрим на следующей лекции
  - ✓ Подход работает только для линейной регрессии
- ✓ Использование методов численной оптимизации, например, метода градиентного спуска
  - ✓ Рассмотрим на этой лекции. Более конкретно будем рассматривать групповой градиентный спуск («Batch» Gradient Descent). Групповой означает, что на каждом этапе градиентного спуска используются все тренировочные примеры
  - ✓ Подход может быть использован и для других методов машинного обучения, например, нейронных сетей

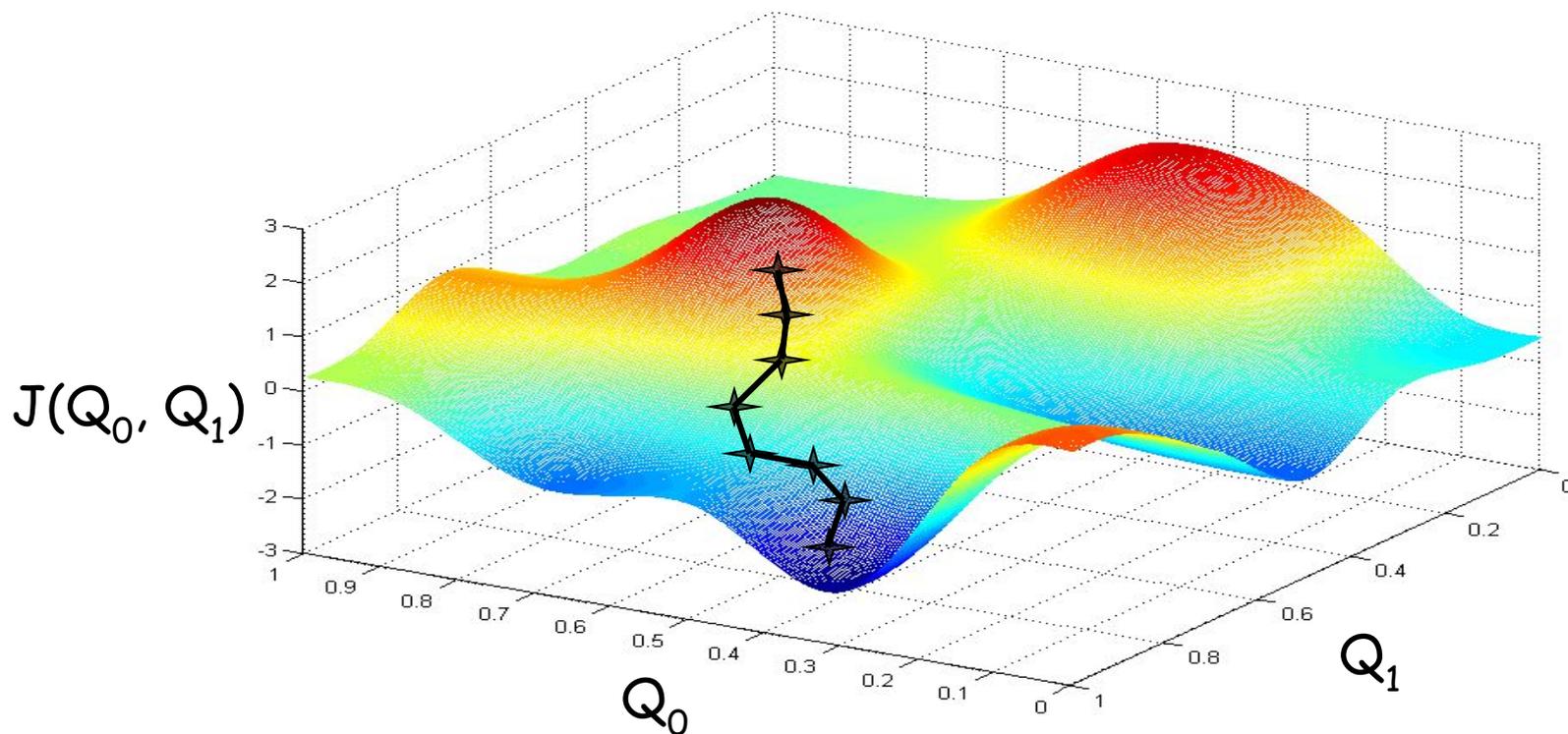


# Метод градиентного спуска

- ✓ Постановка задачи
  - ✓ Имеется некоторая стоимостная функция  $J(Q_0, Q_1)$
  - ✓ Необходимо найти такие значения  $Q_0, Q_1$ , чтобы функция  $J(Q_0, Q_1)$  стала минимальной
- ✓ Решение задачи
  - ✓ Стартуем из некоторых значений  $Q_0, Q_1$ , например, равных величине ноль
  - ✓ Продолжаем изменение значений  $Q_0, Q_1$  до тех пор, пока не достигнем минимума. Минимум достигим не всегда!

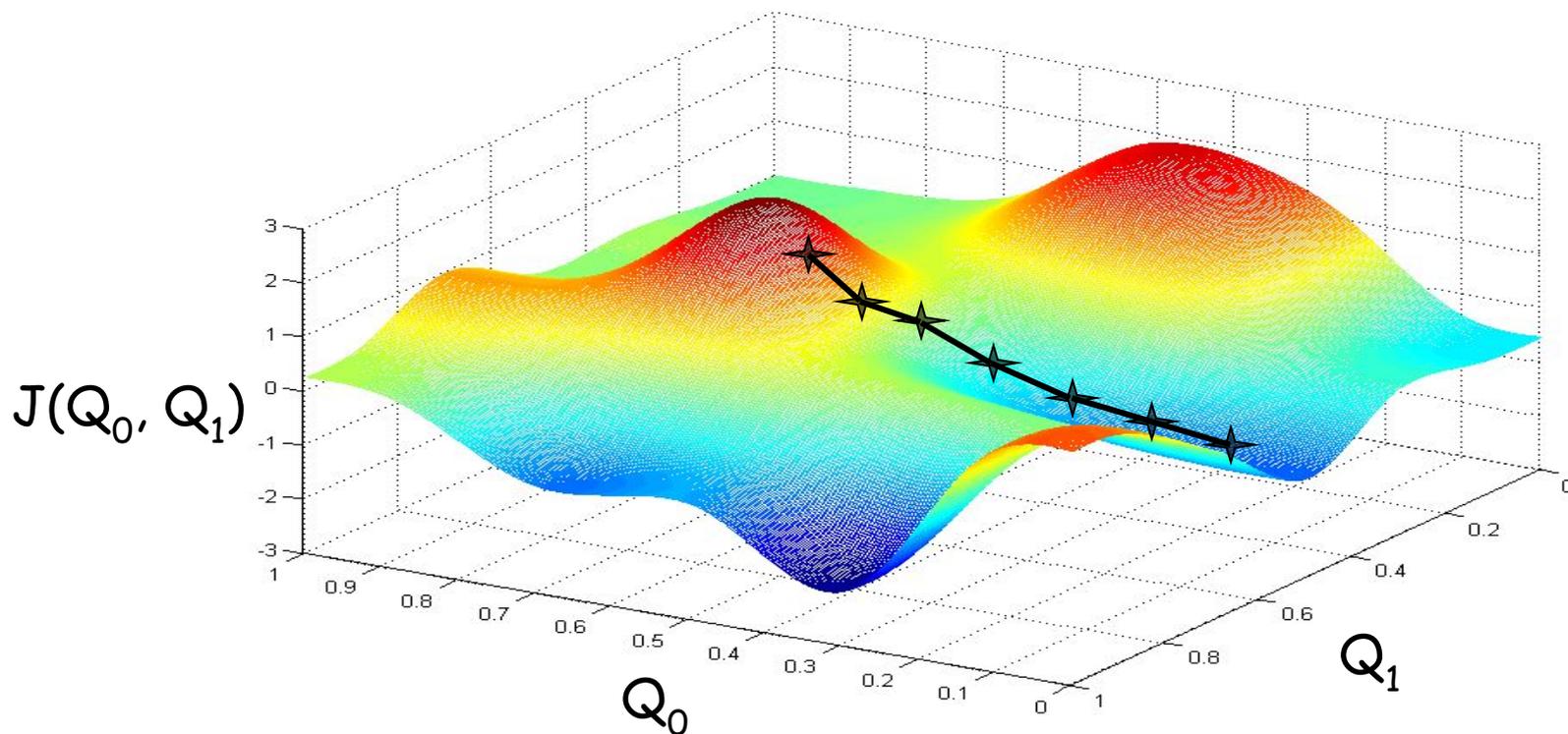


# Пример работы градиентного спуска





# Пример работы градиентного спуска





# Метод градиентного спуска. Реализация

repeat until convergence

$$\left\{ \begin{array}{l} Q_j = Q_j - \alpha \frac{\partial}{\partial Q_j} J(Q_0, Q_1); \quad (j = 0 \text{ и } j = 1) \end{array} \right\}$$

Правильная реализация  
(параметры  $Q_0$ ,  $Q_1$   
обновляются одновременно)

$$\text{temp0} = Q_0 - \alpha \frac{\partial}{\partial Q_0} J(Q_0, Q_1);$$

$$\text{temp1} = Q_1 - \alpha \frac{\partial}{\partial Q_1} J(Q_0, Q_1);$$

$$Q_0 = \text{temp0};$$

$$Q_1 = \text{temp1};$$

Неправильная реализация  
(параметры  $Q_0$ ,  $Q_1$  обновляются  
не одновременно)

$$\text{temp0} = Q_0 - \alpha \frac{\partial}{\partial Q_0} J(Q_0, Q_1);$$

$$Q_0 = \text{temp0};$$

$$\text{temp1} = Q_1 - \alpha \frac{\partial}{\partial Q_1} J(Q_0, Q_1);$$

$$Q_1 = \text{temp1};$$



# Метод градиентного спуска. Реализация

- ✓ Скорость сходимости алгоритма регулируется параметром  $\alpha$ 
  - ✓ Если  $\alpha$  маленькое, то градиентный спуск может быть медленным
  - ✓ Если  $\alpha$  большое, то градиентный спуск может проскочить минимум. Алгоритм может не сходиться или даже расходиться
- ✓ Градиентный спуск может сходиться к локальному минимуму, даже если  $\alpha$  является фиксированным
  - ✓ При приближении к локальному минимуму градиентный спуск будет автоматически выполнять более малые шаги. Поэтому нет необходимости уменьшать  $\alpha$  через некоторое время



# Градиентный спуск для линейной регрессии

repeat until convergence

$$\left\{ \begin{array}{l} Q_j = Q_j - \alpha \frac{\partial}{\partial Q_j} J(Q_0, Q_1); \quad (j = 0 \text{ и } j = 1) \end{array} \right.$$

✓ Вычислив производные получим

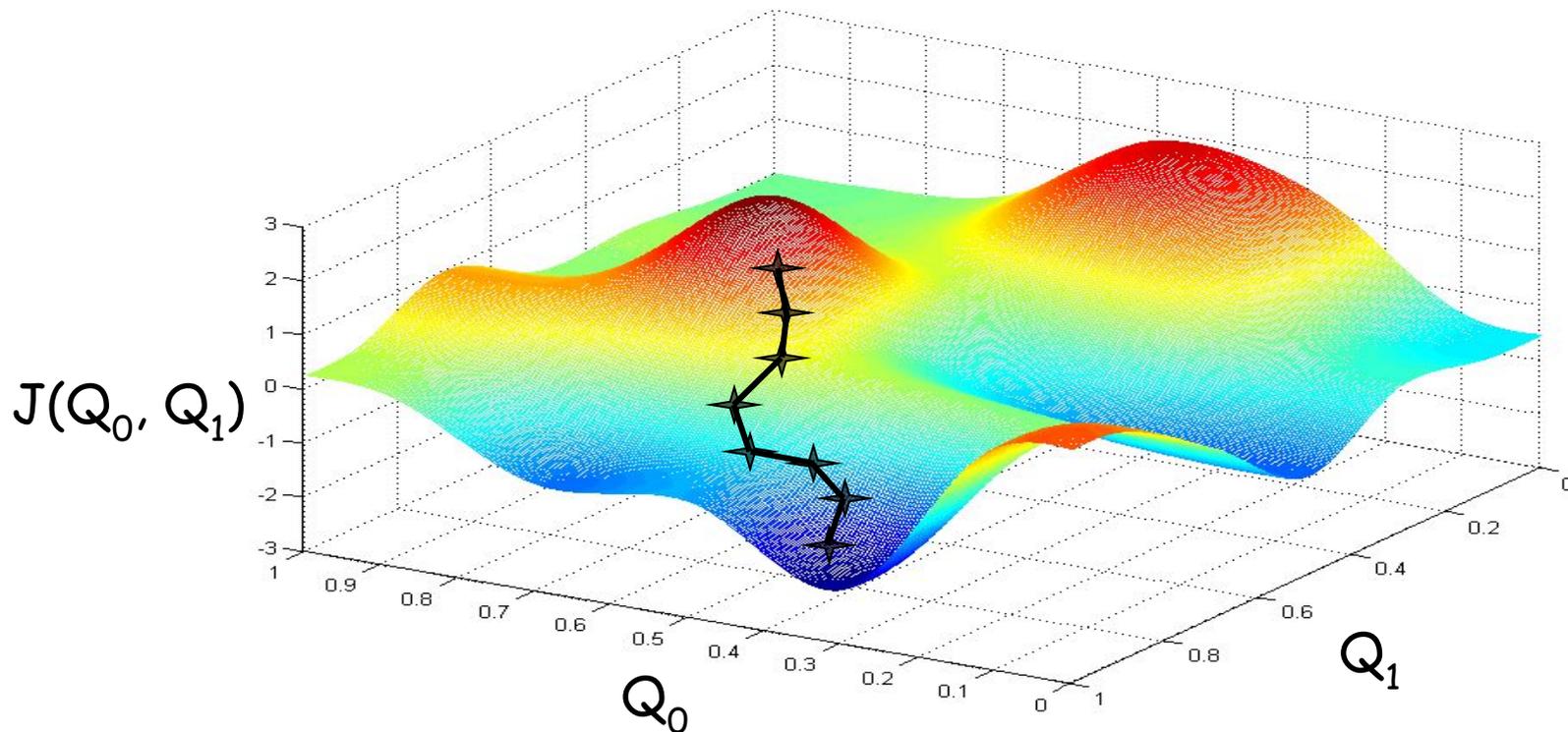
repeat until convergence

$$\left\{ \begin{array}{l} Q_0 = Q_0 - \alpha \frac{1}{m} \sum_{i=1}^m (h_Q(x^{(i)}) - y^{(i)}); \\ Q_1 = Q_1 - \alpha \frac{1}{m} \sum_{i=1}^m (h_Q(x^{(i)}) - y^{(i)}) x^{(i)}; \end{array} \right.$$

параметры  $Q_0$ ,  $Q_1$   
обновляются  
одновременно

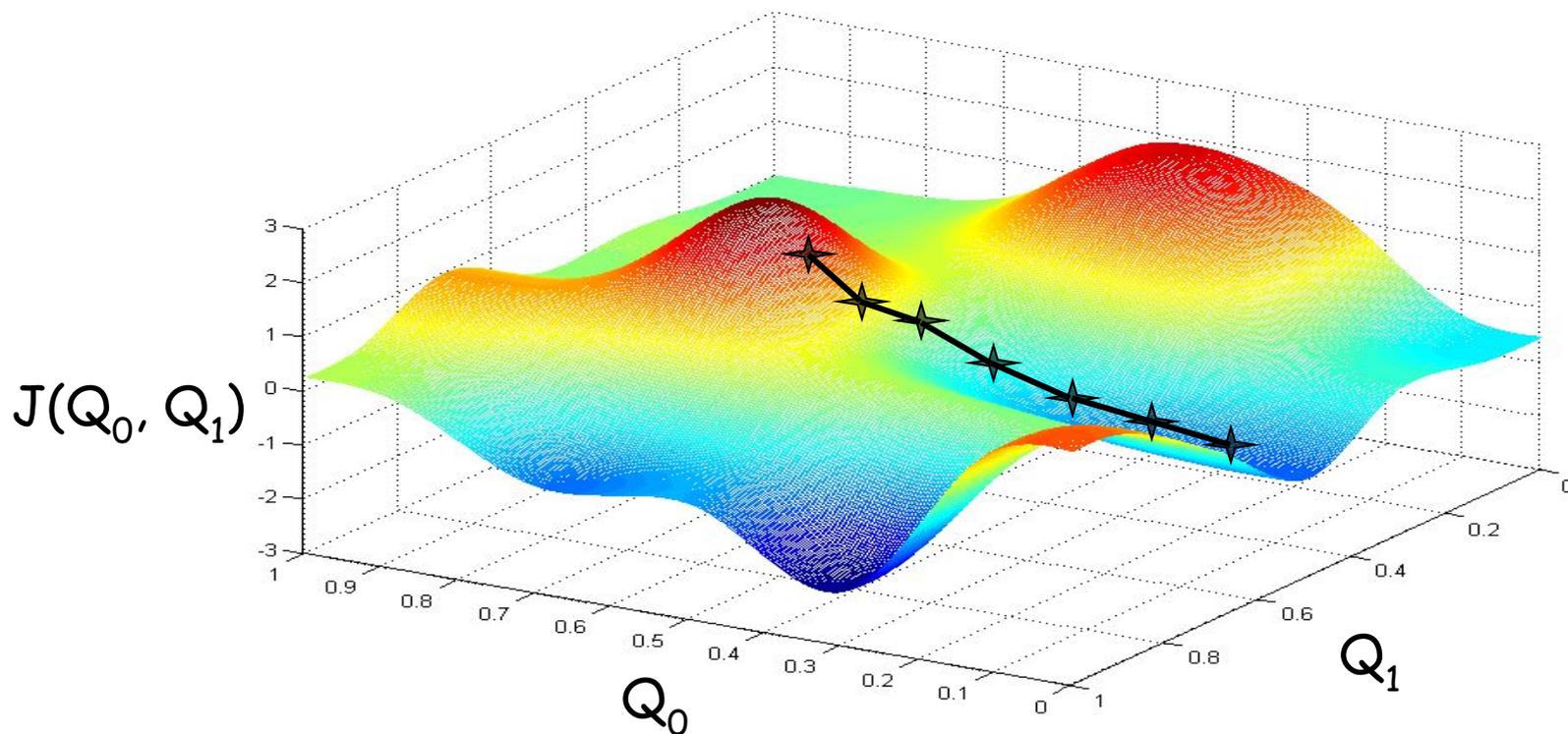


# Пример формы стоимостной функции



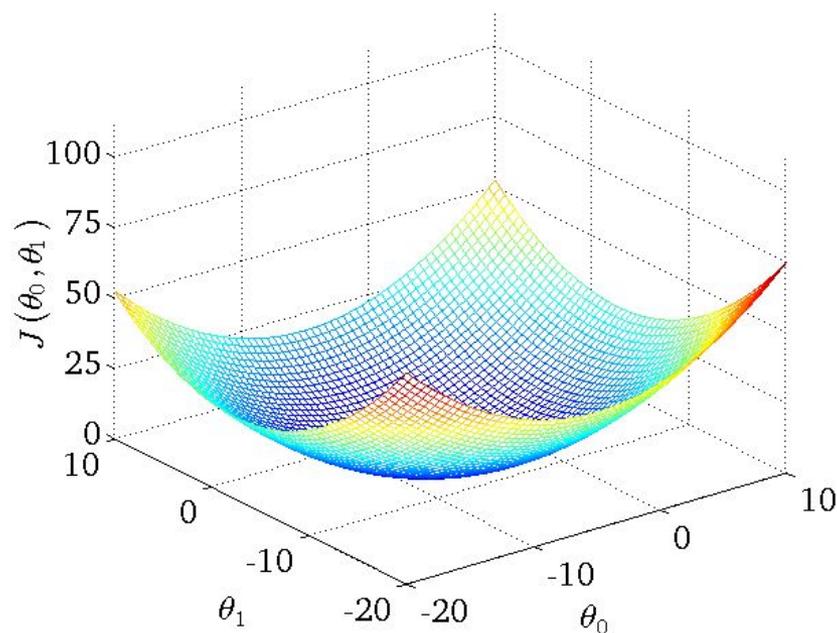


# Пример формы стоимостной функции





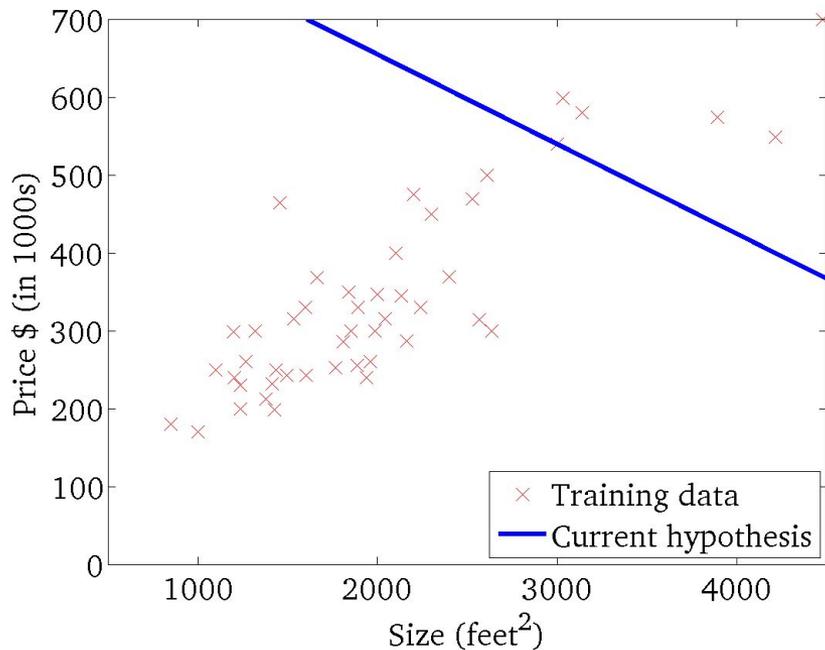
# Пример формы стоимостной функции для линейной регрессии



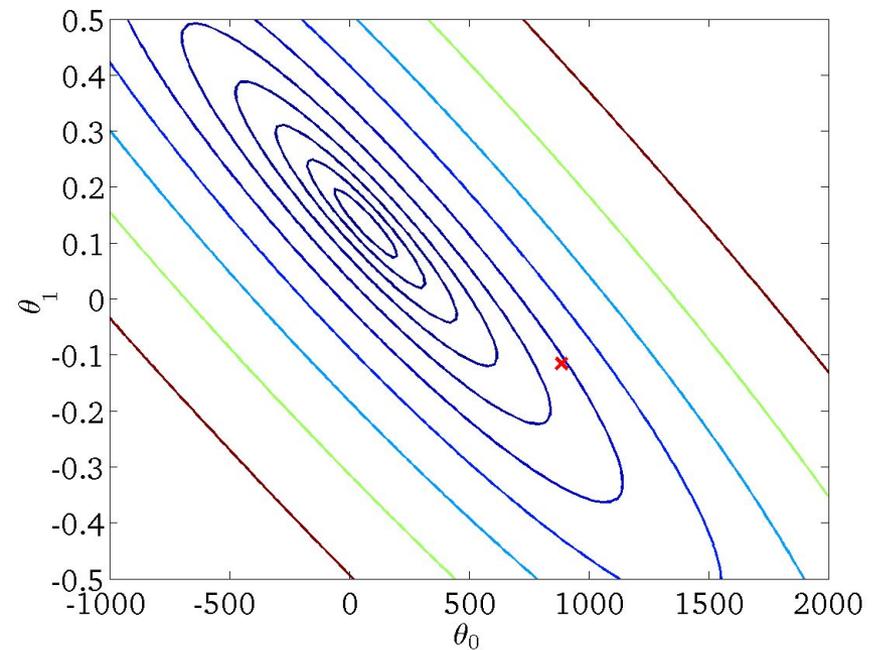
Функция  $J(\theta_0, \theta_1)$  является выпуклой



# Пример работы градиентного спуска для линейной регрессии



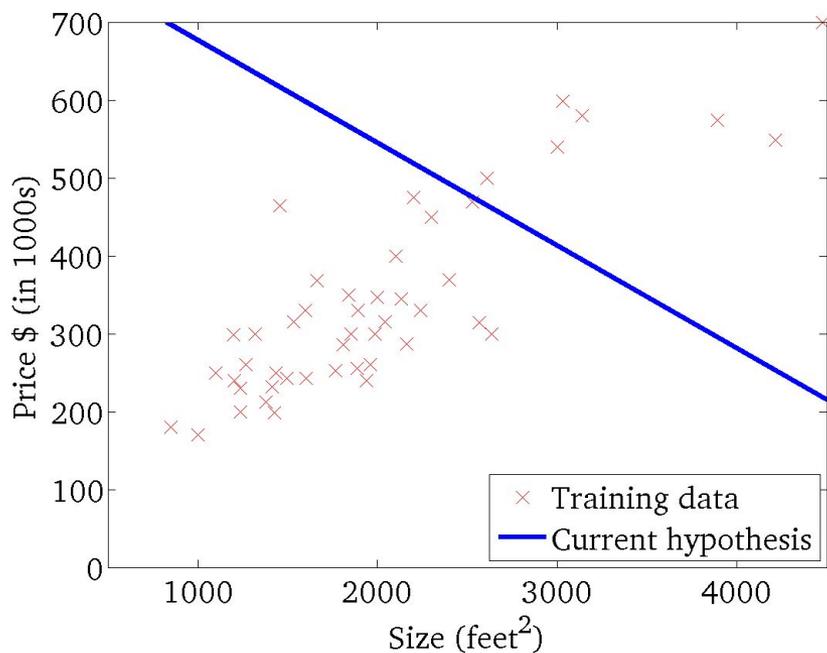
$h_Q(x)$  - это функция  $x$  для фиксированных  $Q_0$  и  $Q_1$



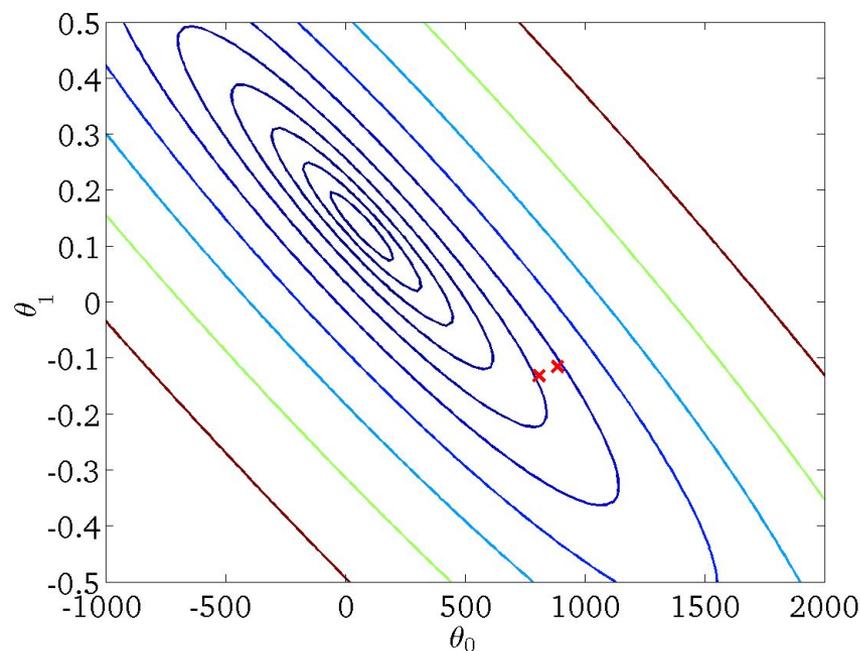
$J(Q_0, Q_1)$  - это функция параметров  $Q_0$  и  $Q_1$



# Пример работы градиентного спуска для линейной регрессии



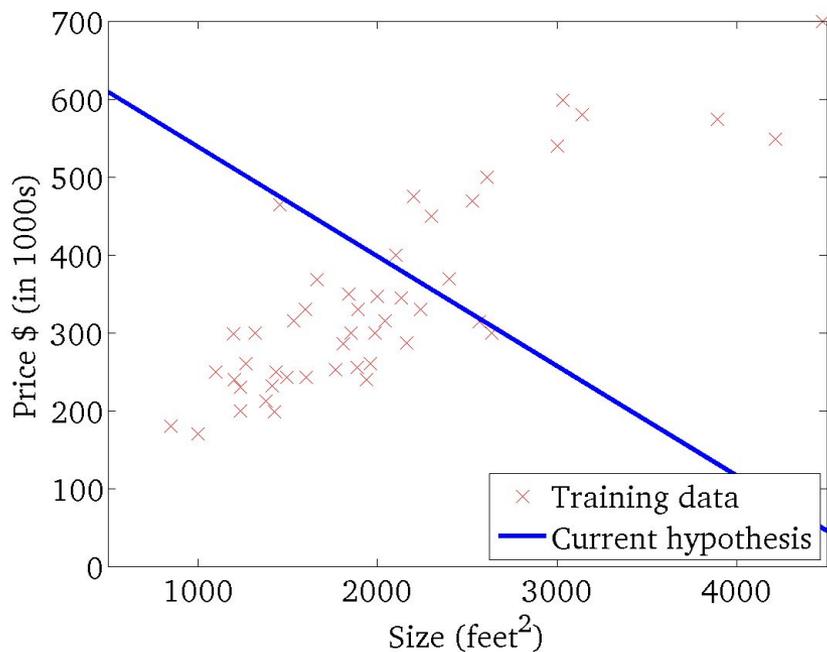
$h_Q(x)$  - это функция  $x$  для фиксированных  $Q_0$  и  $Q_1$



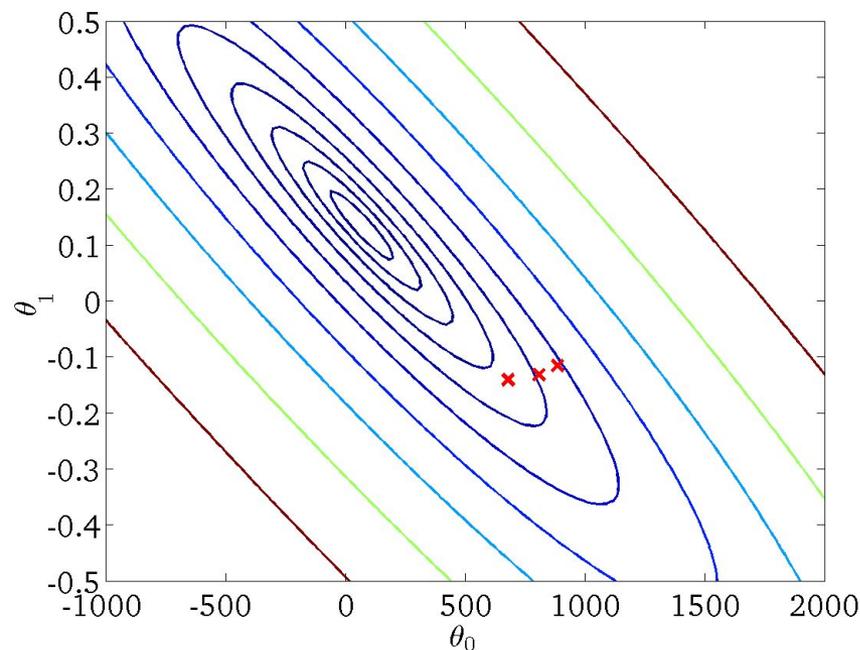
$J(Q_0, Q_1)$  - это функция параметров  $Q_0$  и  $Q_1$



# Пример работы градиентного спуска для линейной регрессии



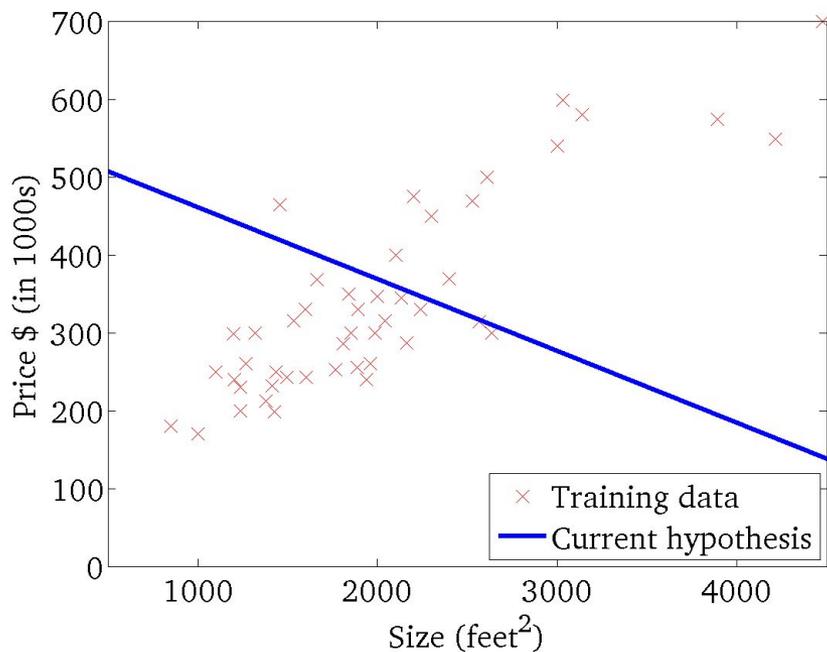
$h_Q(x)$  - это функция  $x$  для фиксированных  $Q_0$  и  $Q_1$



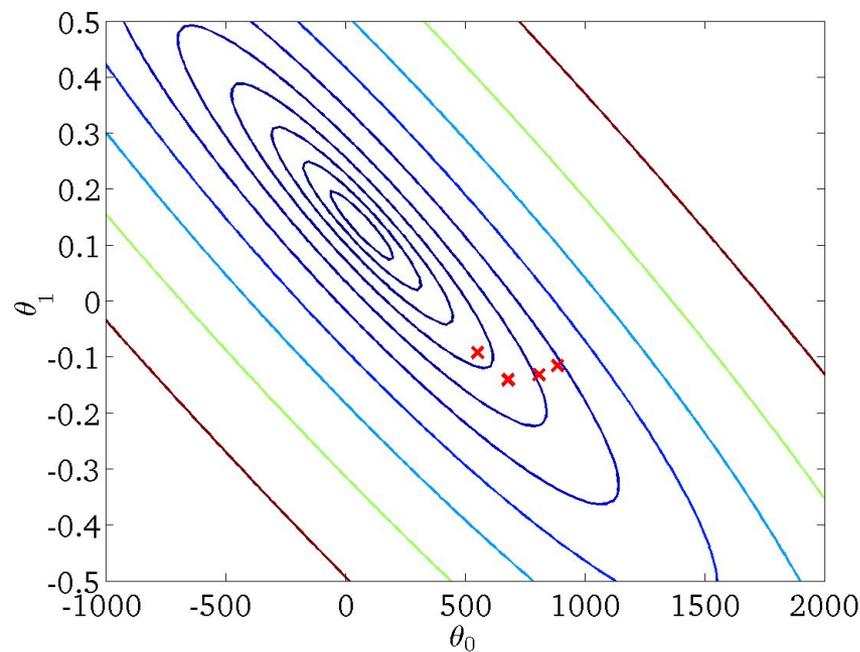
$J(Q_0, Q_1)$  - это функция параметров  $Q_0$  и  $Q_1$



# Пример работы градиентного спуска для линейной регрессии



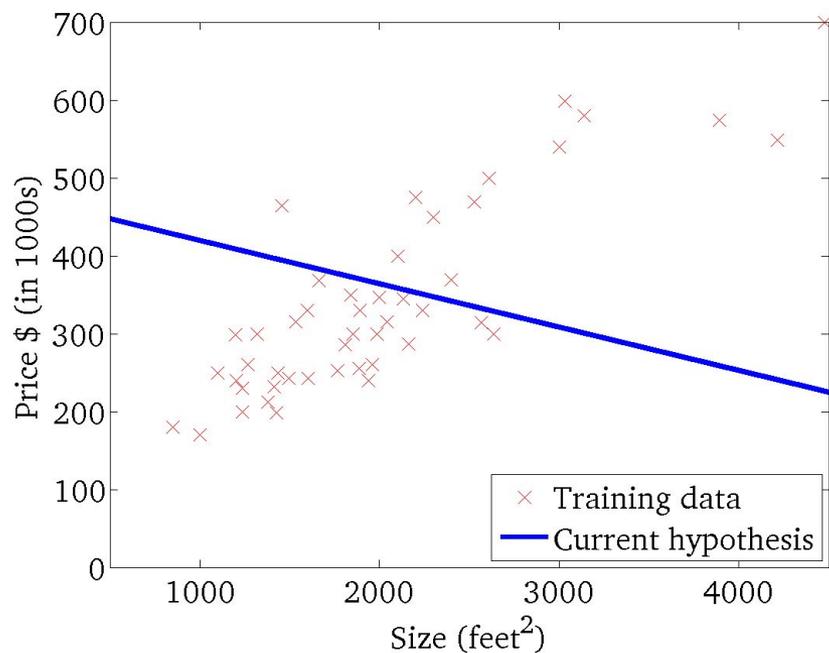
$h_Q(x)$  - это функция  $x$  для фиксированных  $Q_0$  и  $Q_1$



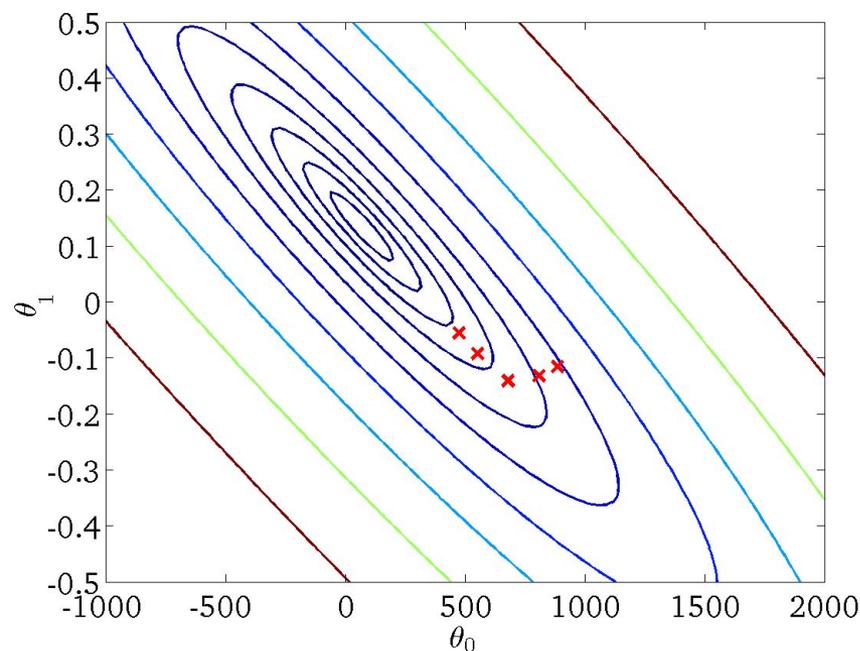
$J(Q_0, Q_1)$  - это функция параметров  $Q_0$  и  $Q_1$



# Пример работы градиентного спуска для линейной регрессии



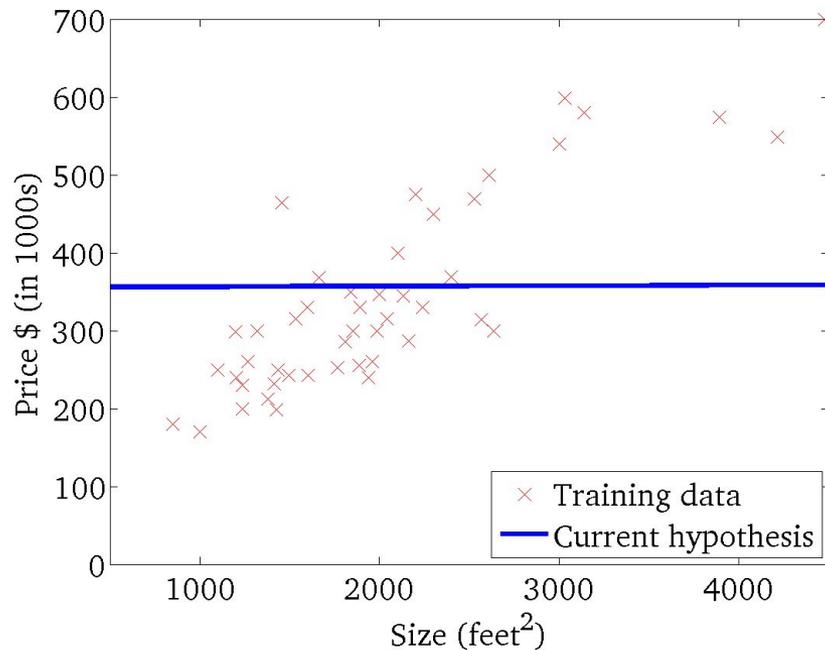
$h_Q(x)$  - это функция  $x$  для фиксированных  $Q_0$  и  $Q_1$



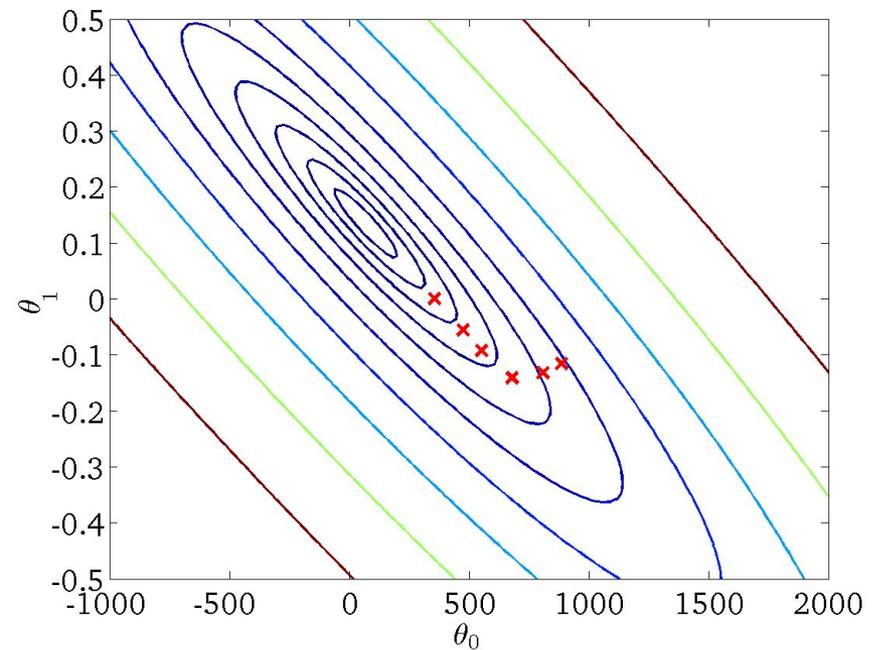
$J(Q_0, Q_1)$  - это функция параметров  $Q_0$  и  $Q_1$



# Пример работы градиентного спуска для линейной регрессии



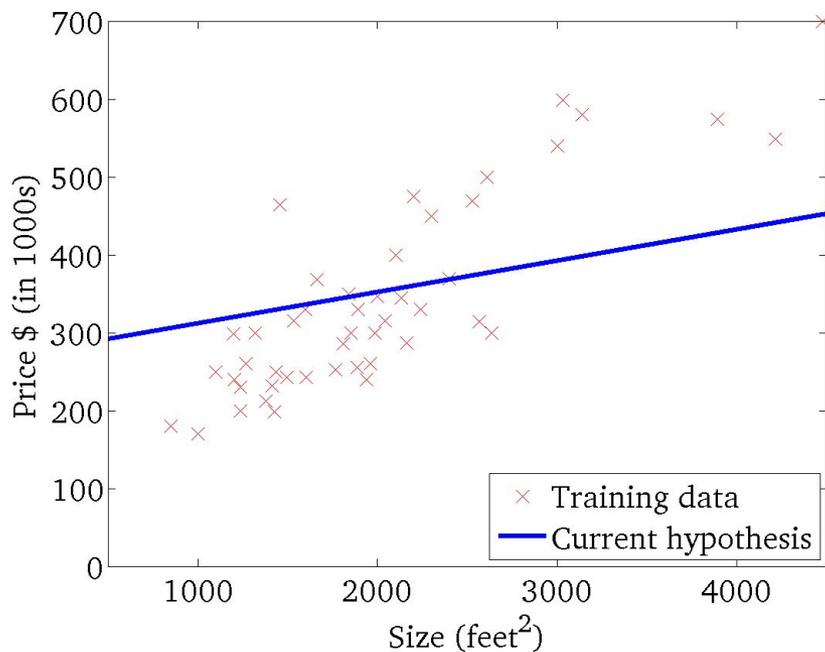
$h_Q(x)$  - это функция  $x$  для фиксированных  $Q_0$  и  $Q_1$



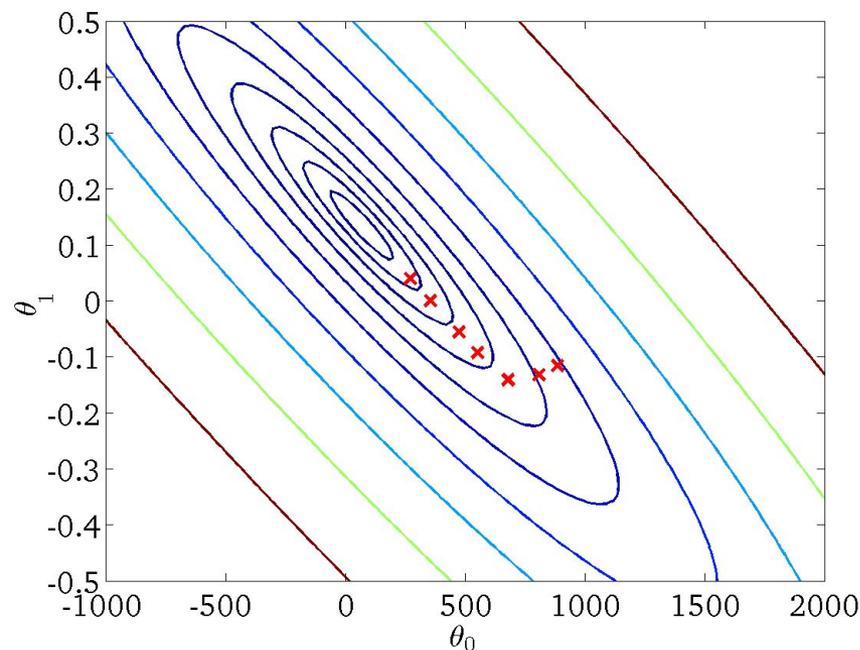
$J(Q_0, Q_1)$  - это функция параметров  $Q_0$  и  $Q_1$



# Пример работы градиентного спуска для линейной регрессии



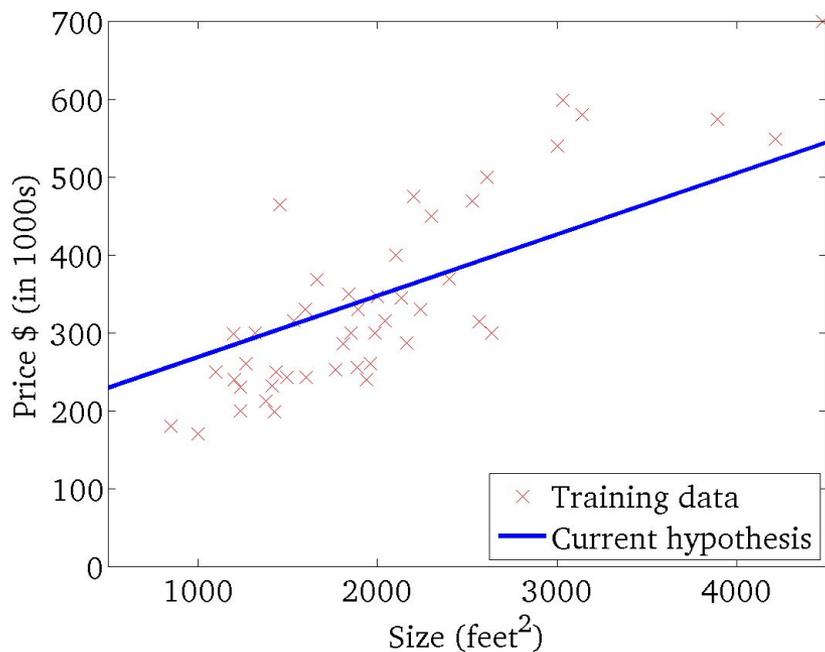
$h_Q(x)$  - это функция  $x$  для фиксированных  $Q_0$  и  $Q_1$



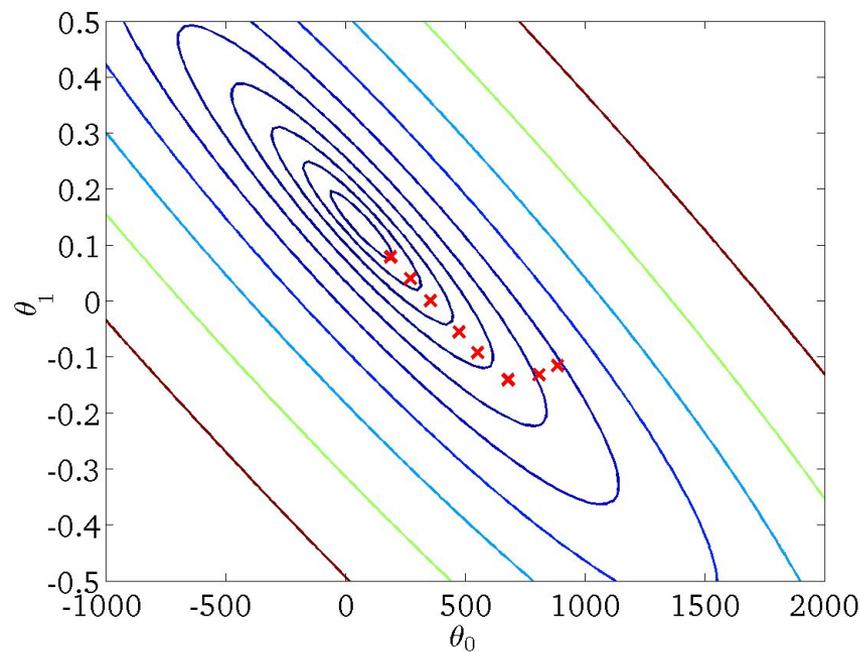
$J(Q_0, Q_1)$  - это функция параметров  $Q_0$  и  $Q_1$



# Пример работы градиентного спуска для линейной регрессии



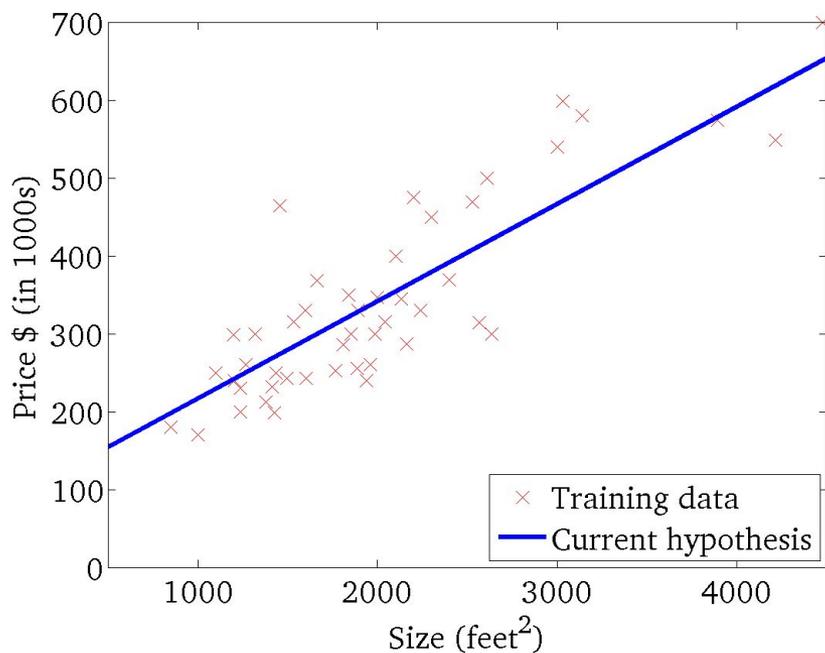
$h_Q(x)$  - это функция  $x$  для фиксированных  $Q_0$  и  $Q_1$



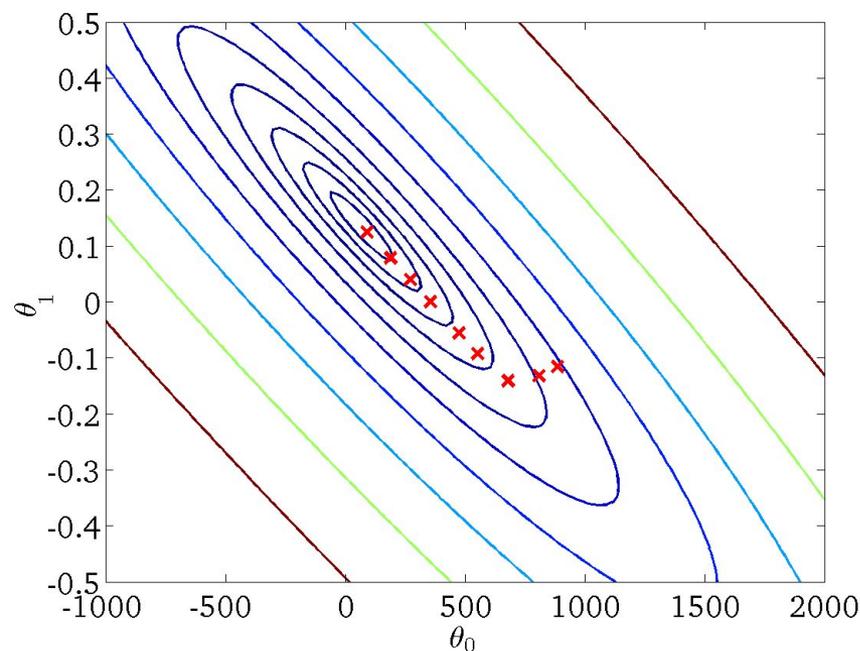
$J(Q_0, Q_1)$  - это функция параметров  $Q_0$  и  $Q_1$



# Пример работы градиентного спуска для линейной регрессии



$h_Q(x)$  - это функция  $x$  для фиксированных  $Q_0$  и  $Q_1$



$J(Q_0, Q_1)$  - это функция параметров  $Q_0$  и  $Q_1$



## Данные, данные, данные ...

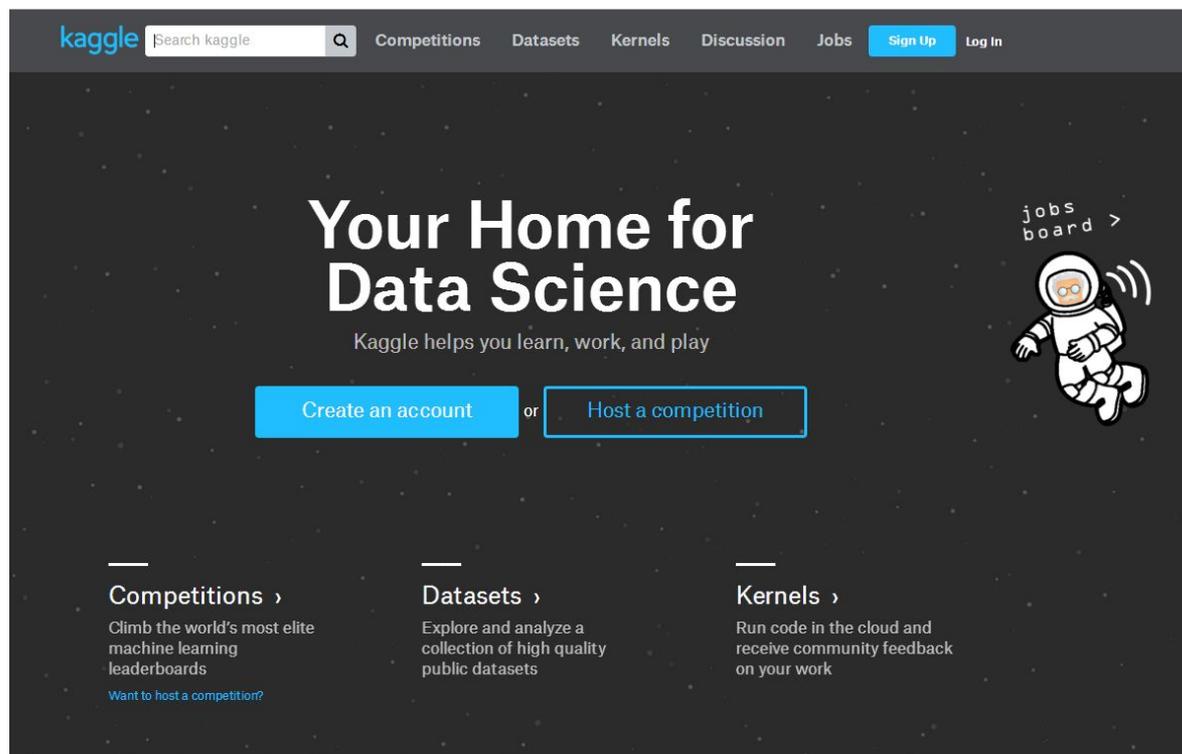
- ✓ К.В. Воронцов (ШАД Яндекс): «Кто обучает машину? Может быть мы? Может быть вы? Оба ответа неверны. Машина обучается на данных! Они называются обучающей выборкой»





# Kaggle помогает вам учиться, работать и играть

✓ Соревнования, базы данных и программные коды на [www.kaggle.com](http://www.kaggle.com)





# Конкурс PASCAL VOC

- ✓ Pattern Analysis, Statistical Modeling and Computational Learning Visual Object Classes Challenge
  - ✓ Конкурс больше не проводится
  - ✓ Доступны материалы прошлых лет (базы данных, отчеты и т.п.)





# База данных ImageNet

✓ Большая база данных цифровых изображений

IMAGENET

14,197,122 images, 21841 synsets indexed

[Explore](#) [Download](#) [Challenges](#) [Publications](#) [CoolStuff](#) [About](#)

Not logged in. [Login](#) | [Signup](#)

**ImageNet** is an image database organized according to the **WordNet** hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images. Currently we have an average of over five hundred images per node. We hope ImageNet will become a useful resource for researchers, educators, students and all of you who share our passion for pictures.

[Click here](#) to learn more about ImageNet, [Click here](#) to join the ImageNet mailing list.



What do these images have in common? *Find out!*

[Check out the ImageNet Challenge 2016](#)

© 2016 Stanford Vision Lab, Stanford University, Princeton University support@image-net.org Copyright infringement



# Благодарности

- ✓ В лекции использовались материалы курса:
- ✓ Andrew Ng. Machine Learning (online class), 2012. Stanford University, [www.coursera.org/course/ml](http://www.coursera.org/course/ml)



Куррикулум витте Эндрю здесь: <http://ai.stanford.edu/~ang/>