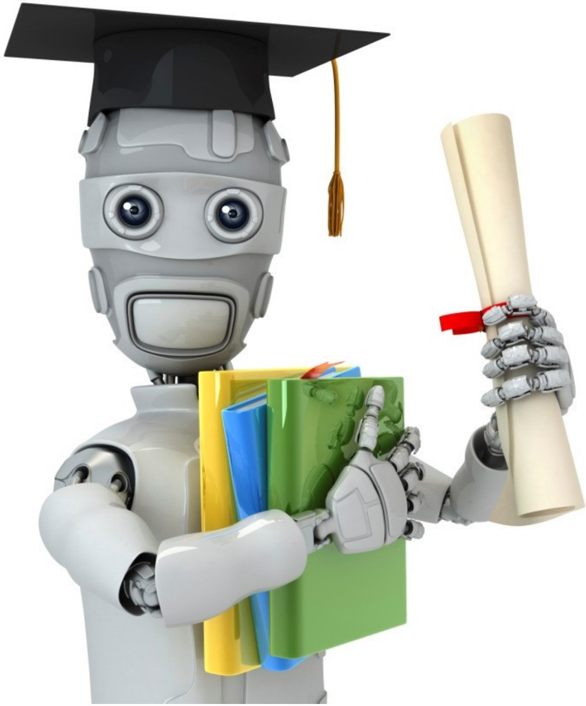
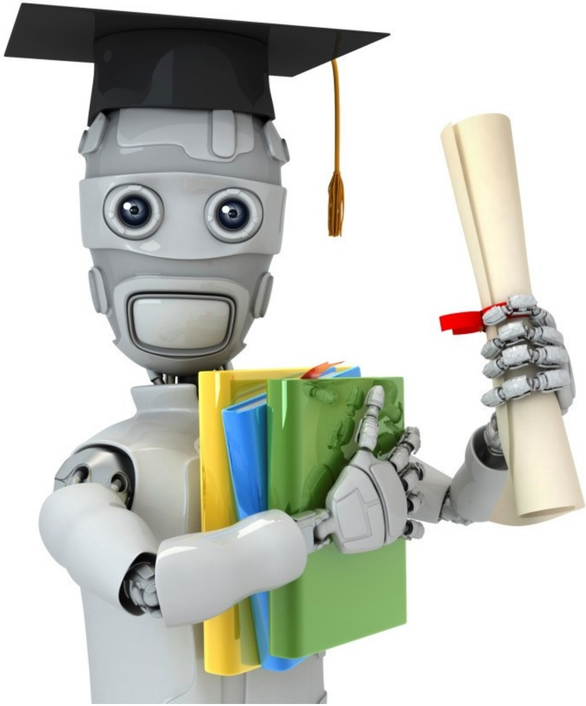

Machine Learning

Introduction

Welcome
Sachin Joshi –
Research Computing



Machine Learning

Contents


Apple - iPhoto - New full-s
www.apple.com/ilife/iphoto/
Store
Mac
iPod
iPhone
iPad
iTunes
Support
iLife '11
iPhoto
iMovie
GarageBand
Video Showcase
Resources
Upgrade Now
iPhoto '11
From your Facebook Wall to your coffee table to your best friend's inbox (or mailbox). Do more with your photos than you ever thought possible. And do it all in one place. iPhoto.
Watch the iPhoto video
MacBook Pro
What's New in iPhoto
What is iPhoto?


SPAM

Machine Learning

- Grew out of work in AI
- New capability for computers

Examples:

- Database mining

  Large datasets from growth of automation/web.

  E.g., Web click data, medical records, biology, engineering
- Applications can’t program by hand.

  E.g., Autonomous helicopter, handwriting recognition, most of Natural Language Processing (NLP), Computer Vision.

Machine Learning

- Grew out of work in AI
- New capability for computers

Examples:

- Database mining

  Large datasets from growth of automation/web.

  E.g., Web click data, medical records, biology, engineering
- Applications can't program by hand.

  E.g., Autonomous helicopter, handwriting recognition, most of Natural Language Processing (NLP), Computer Vision.
- Self-customizing programs

  E.g., Amazon, Netflix product recommendations

Machine Learning

- Grew out of work in AI
- New capability for computers

Examples:

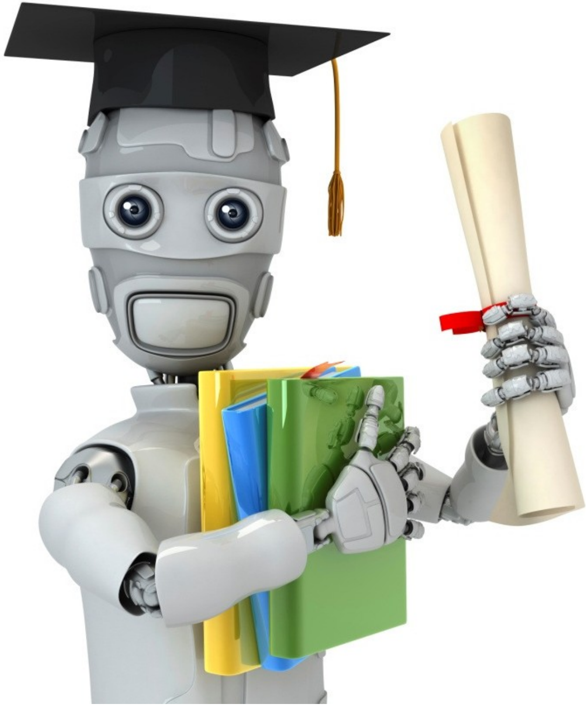
- Database mining

  Large datasets from growth of automation/web.

  E.g., Web click data, medical records, biology, engineering
- Applications can’t program by hand.

  E.g., Autonomous helicopter, handwriting recognition, most of Natural Language Processing (NLP), Computer Vision.
- Self-customizing programs

  E.g., Amazon, Netflix product recommendations
- Understanding human learning (brain, real AI).



Machine Learning

Introduction

What is machine learning

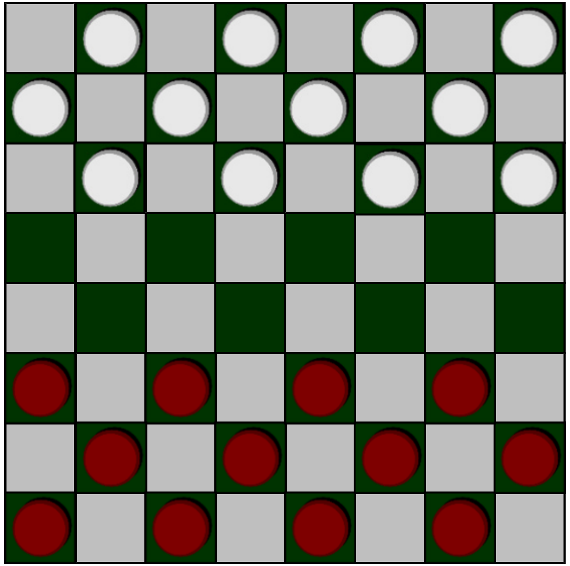
Machine Learning definition

Machine Learning definition

- Arthur Samuel (1959). Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed.

Machine Learning definition

- Arthur Samuel (1959). Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed.



Machine Learning definition

- Arthur Samuel (1959). Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed.
- Tom Mitchell (1998) Well-posed Learning Problem: A computer program is said to *learn* from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

"A computer program is said to *learn* from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E."

Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is the task T in this setting?

- ○ Classifying emails as spam or not spam.
- ○ Watching you label emails as spam or not spam.
- ○ The number (or fraction) of emails correctly classified as spam/not spam.
- ○ None of the above—this is not a machine learning problem.

"A computer program is said to *learn* from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E."

Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is the task T in this setting?

- ◯ Classifying emails as spam or not spam. T
- ◯ Watching you label emails as spam or not spam. E
- ◯ The number (or fraction) of emails correctly classified as spam/not spam. P
- ◯ None of the above—this is not a machine learning problem.

“A computer program is said to *learn* from experience E with respect to

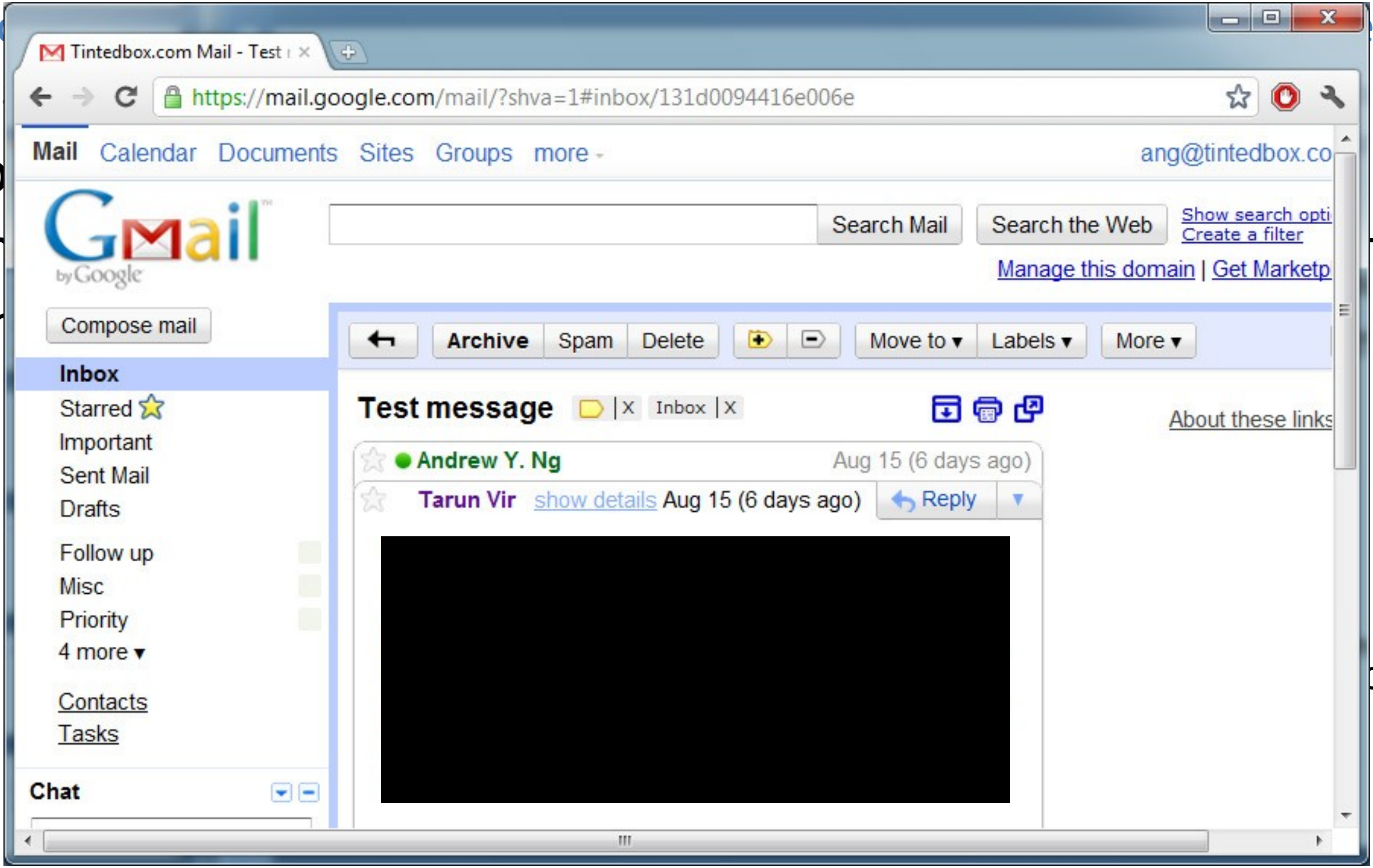
s … on T,

a

Supp … or do

not r … filter

spam



pam.

"A computer program is said to *learn* from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E."

Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam. What is the task T in this setting?

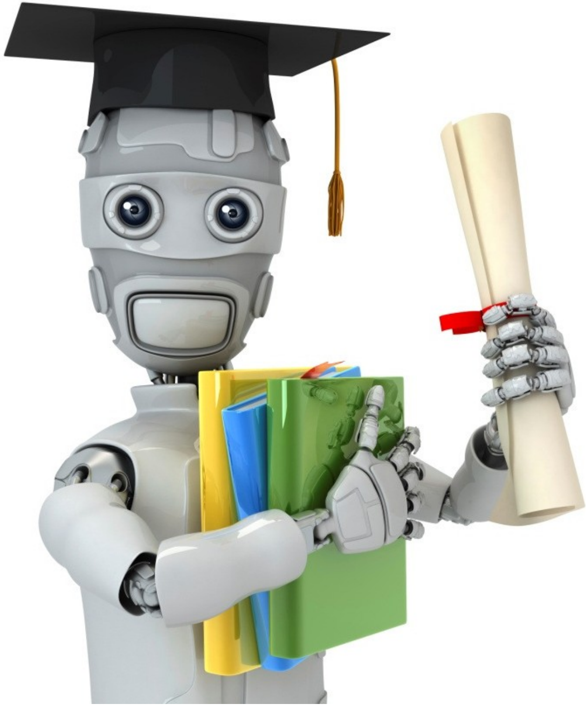
- ◯ Classifying emails as spam or not spam. T
- ◯ Watching you label emails as spam or not spam. E
- ◯ The number (or fraction) of emails correctly classified as spam/not spam. P
- ◯ None of the above—this is not a machine learning problem.

Machine learning algorithms:

- Supervised learning
- Unsupervised learning

Others: Reinforcement learning, recommender systems.

Also talk about: Practical advice for applying learning algorithms.

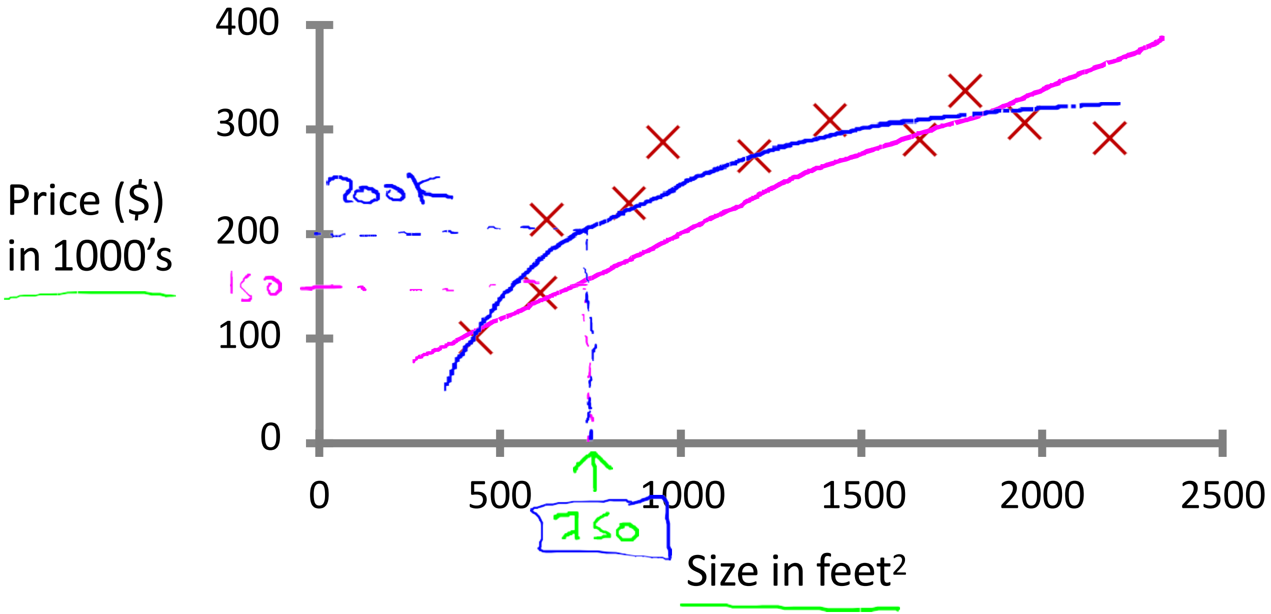


Machine Learning

Introduction

Supervised Learning

Housing price prediction.

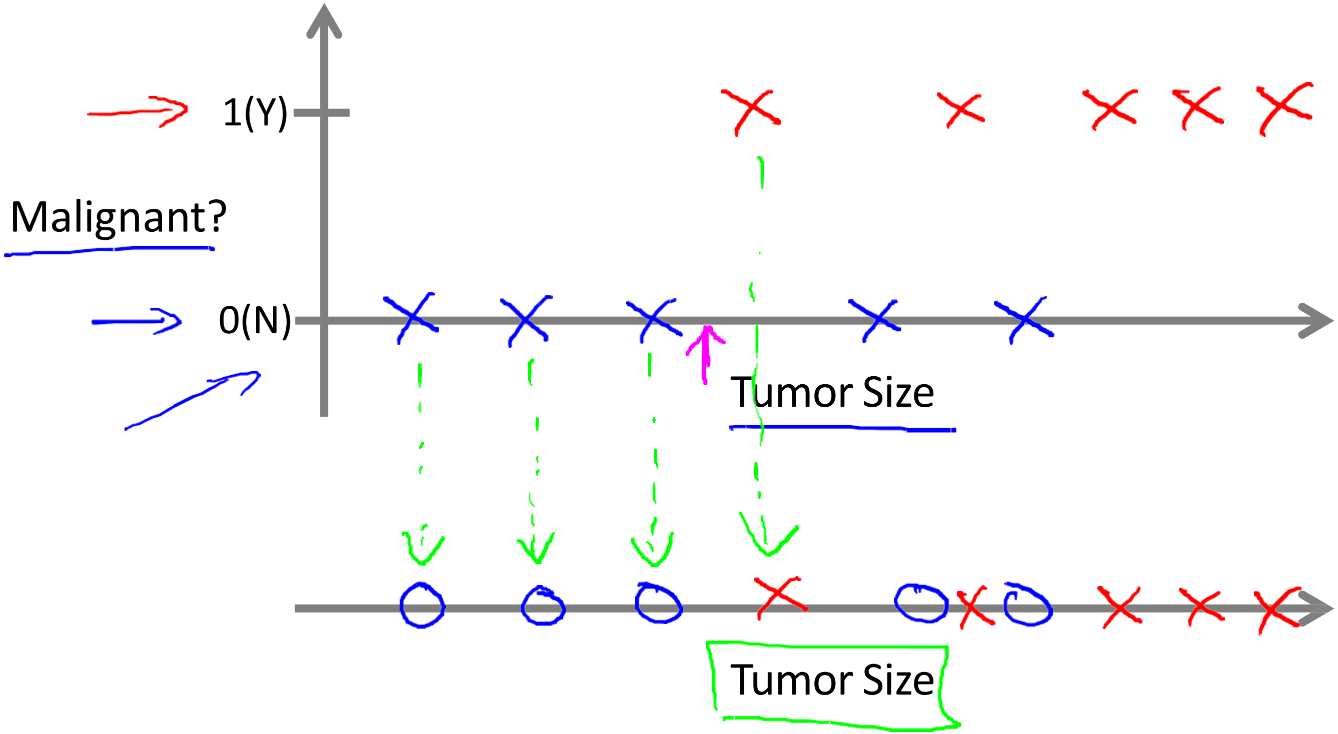


Supervised Learning

"right answers" given

Regression: Predict continuous valued output (price)

Breast cancer (malignant, benign)

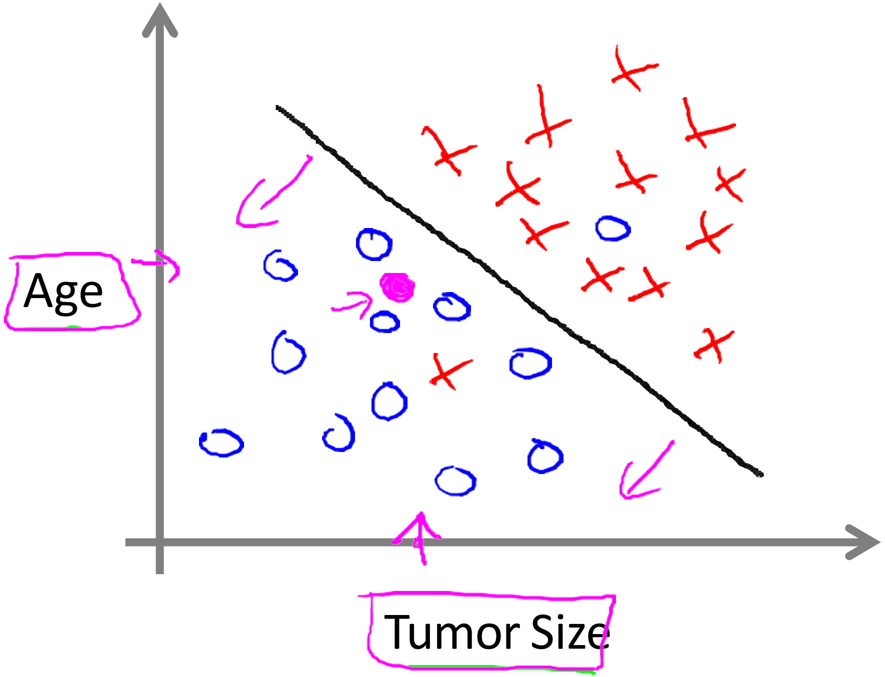


Classification

Discrete valued output (0 or 1)

0, 1, 2, 3

benign type 1 cancer



- Clump Thickness
- Uniformity of Cell Size
- Uniformity of Cell Shape

...

You’re running a company, and you want to develop learning algorithms to address each of two problems.

Problem 1: You have a large inventory of identical items. You want to predict how many of these items will sell over the next 3 months.
Problem 2: You’d like software to examine individual customer accounts, and for each account decide if it has been hacked/compromised.

Should you treat these as classification or as regression problems?

- ○ Treat both as classification problems.
- ○ Treat problem 1 as a classification problem, problem 2 as a regression problem.
- ○ Treat problem 1 as a regression problem, problem 2 as a classification problem.
- ○ Treat both as regression problems.

You're running a company, and you want to develop learning algorithms to address each of two problems.

1000's

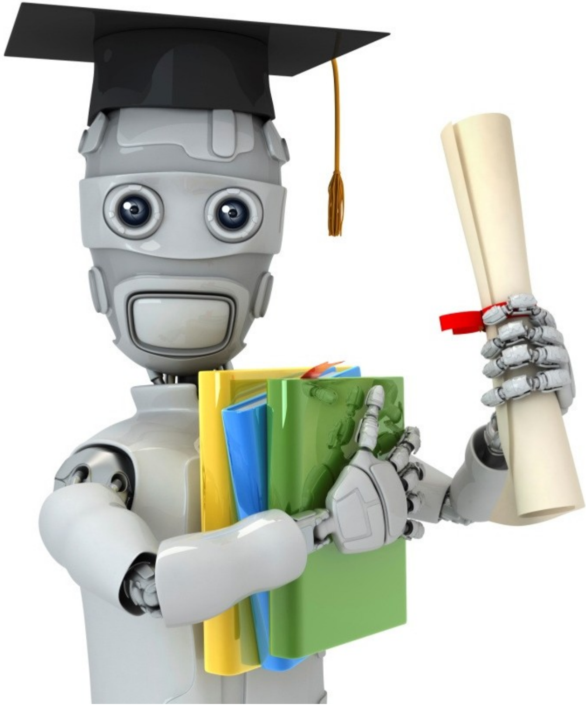
Problem 1: You have a large inventory of identical items. You want to predict how many of these items will sell over the next 3 months.

Problem 2: You'd like software to examine individual customer accounts, and for each account decide if it has been hacked/compromised.

0 – not hacked
1 – hacked

Should you treat these as classification or as regression problems?

- ◯ Treat both as classification problems.
- ◯ Treat problem 1 as a classification problem, problem 2 as a regression problem.
- ◯ Treat problem 1 as a regression problem, problem 2 as a classification problem.
- ◯ Treat both as regression problems.

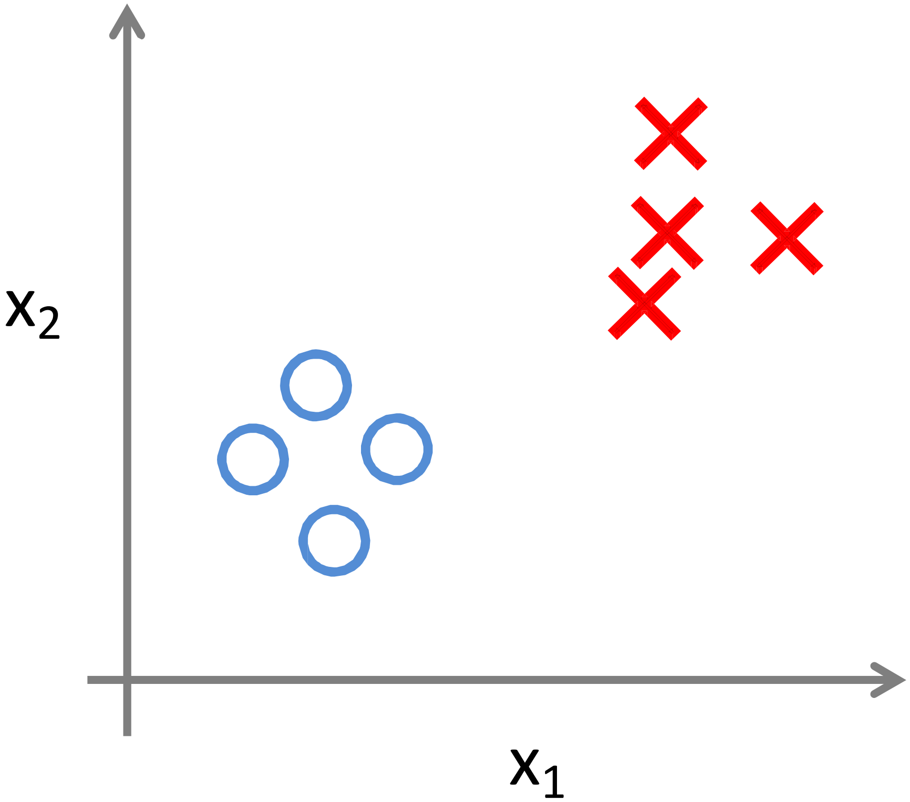


Machine Learning

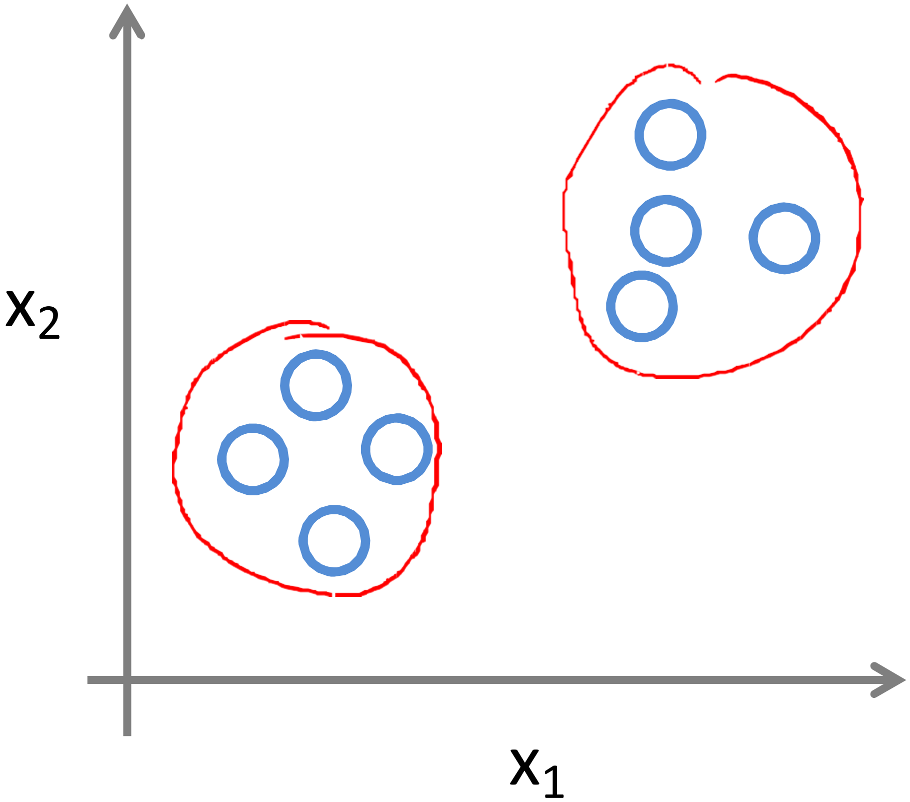
Introduction

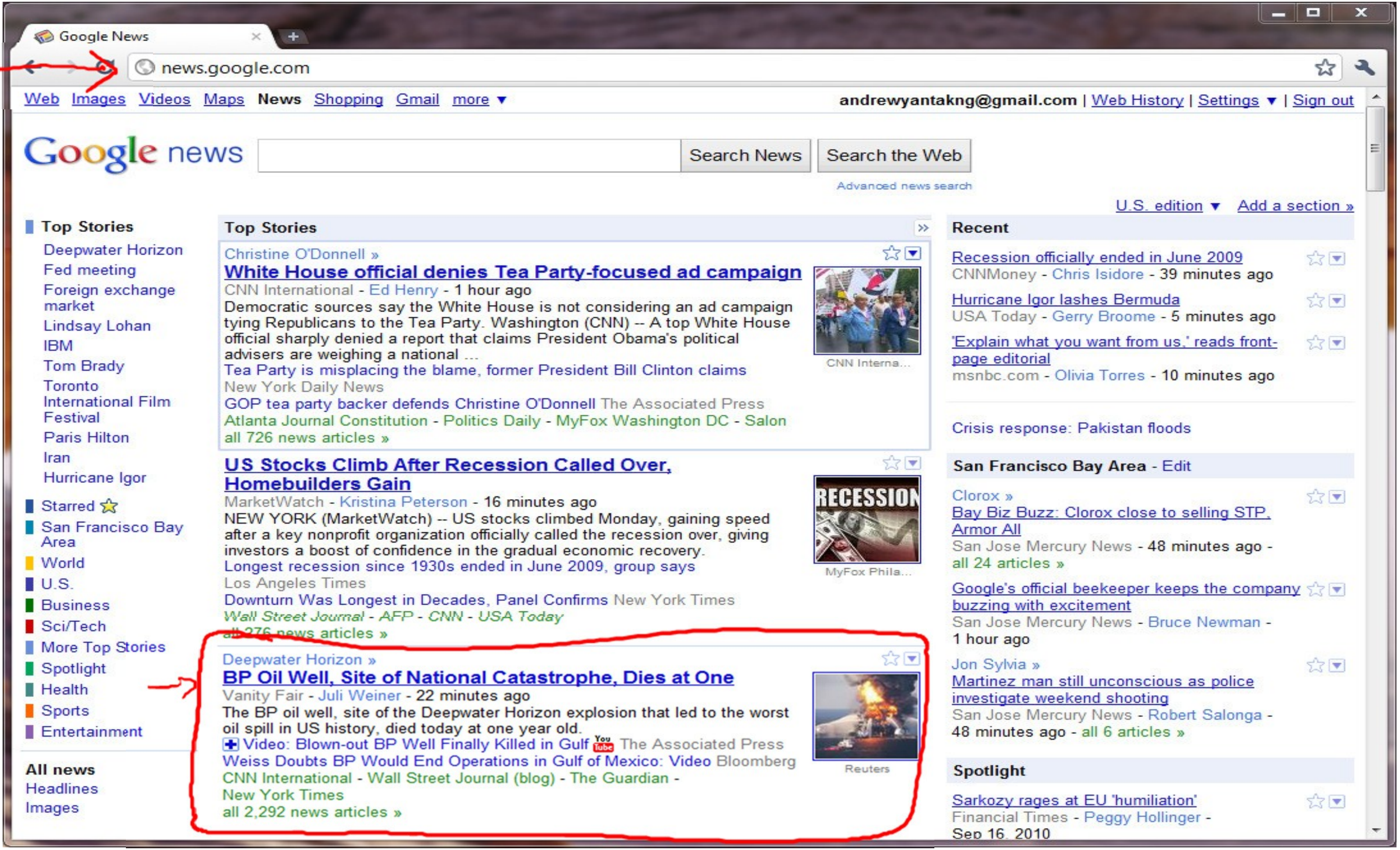
Unsupervised Learning

Supervised Learning



Unsupervised Learning




Google News

news.google.com

Web Images Videos Maps News Shopping Gmail more ▾ andrewyantakng@gmail.com | Web History | Settings ▾ | Sign out

Google news

Search News | Search the Web

Advanced news search

U.S. edition ▾ Add a section »

Top Stories
- Deepwater Horizon
- Fed meeting
- Foreign exchange market
- Lindsay Lohan
- IBM
- Tom Brady
- Toronto International Film Festival
- Paris Hilton
- Iran
- Hurricane Igor

- Starred
- San Francisco Bay Area
- World
- U.S.
- Business
- Sci/Tech
- More Top Stories
- Spotlight
- Health
- Sports
- Entertainment

All news
- Headlines
- Images

Top Stories

Christine O'Donnell »

White House official denies Tea Party-focused ad campaign

CNN International - Ed Henry - 1 hour ago

Democratic sources say the White House is not considering an ad campaign tying Republicans to the Tea Party. Washington (CNN) -- A top White House official sharply denied a report that claims President Obama's political advisers are weighing a national ...

Tea Party is misplacing the blame, former President Bill Clinton claims New York Daily News

GOP tea party backer defends Christine O'Donnell The Associated Press

Atlanta Journal Constitution - Politics Daily - MyFox Washington DC - Salon

all 726 news articles »

CNN Interna...

US Stocks Climb After Recession Called Over, Homebuilders Gain

MarketWatch - Kristina Peterson - 16 minutes ago

NEW YORK (MarketWatch) -- US stocks climbed Monday, gaining speed after a key nonprofit organization officially called the recession over, giving investors a boost of confidence in the gradual economic recovery.

Longest recession since 1930s ended in June 2009, group says Los Angeles Times

Downturn Was Longest in Decades, Panel Confirms New York Times

Wall Street Journal - AFP - CNN - USA Today

all 276 news articles »

RECESSION

MyFox Phila...

Deepwater Horizon »

BP Oil Well, Site of National Catastrophe, Dies at One

Vanity Fair - Juli Weiner - 22 minutes ago

The BP oil well, site of the Deepwater Horizon explosion that led to the worst oil spill in US history, died today at one year old.

Video: Blown-out BP Well Finally Killed in Gulf The Associated Press

Weiss Doubts BP Would End Operations in Gulf of Mexico: Video Bloomberg

CNN International - Wall Street Journal (blog) - The Guardian - New York Times

all 2,292 news articles »

Reuters

Recent

Recession officially ended in June 2009
CNNMoney - Chris Isidore - 39 minutes ago

Hurricane Igor lashes Bermuda
USA Today - Gerry Broome - 5 minutes ago

'Explain what you want from us,' reads front-page editorial
msnbc.com - Olivia Torres - 10 minutes ago

Crisis response: Pakistan floods

San Francisco Bay Area - Edit

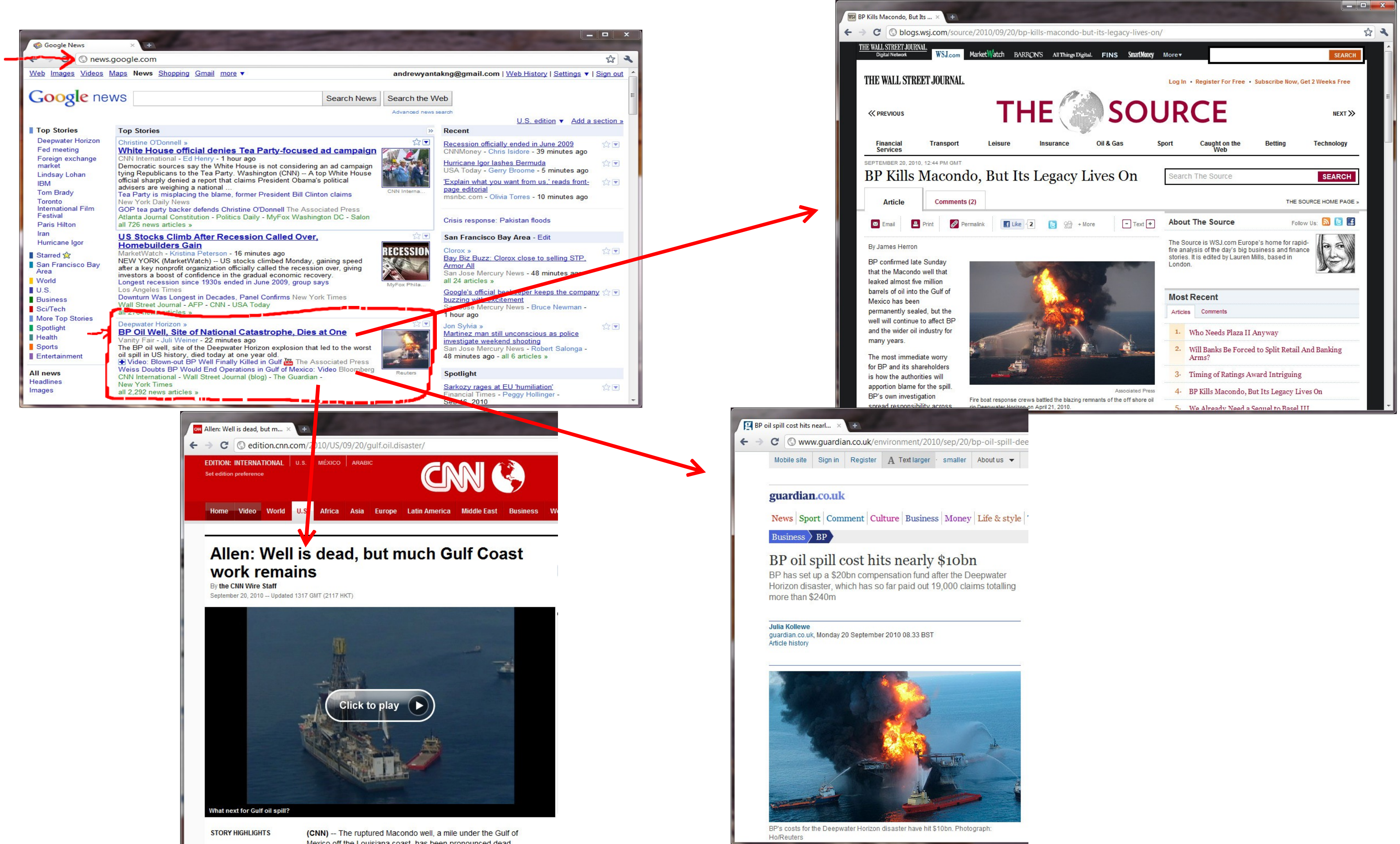
Clorox »
Bay Biz Buzz: Clorox close to selling STP, Armor All
San Jose Mercury News - 48 minutes ago - all 24 articles »

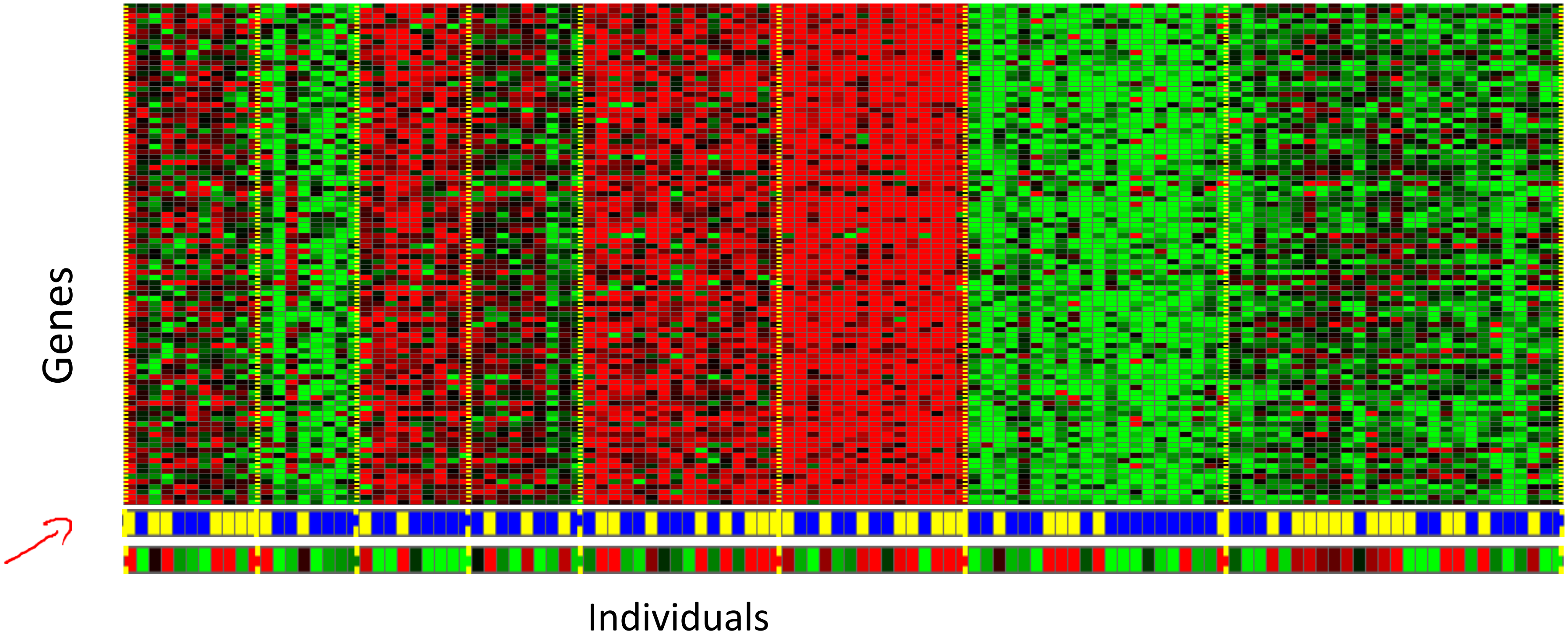
Google's official beekeeper keeps the company buzzing with excitement
San Jose Mercury News - Bruce Newman - 1 hour ago

Jon Sylvia »
Martinez man still unconscious as police investigate weekend shooting
San Jose Mercury News - Robert Salonga - 48 minutes ago - all 6 articles »

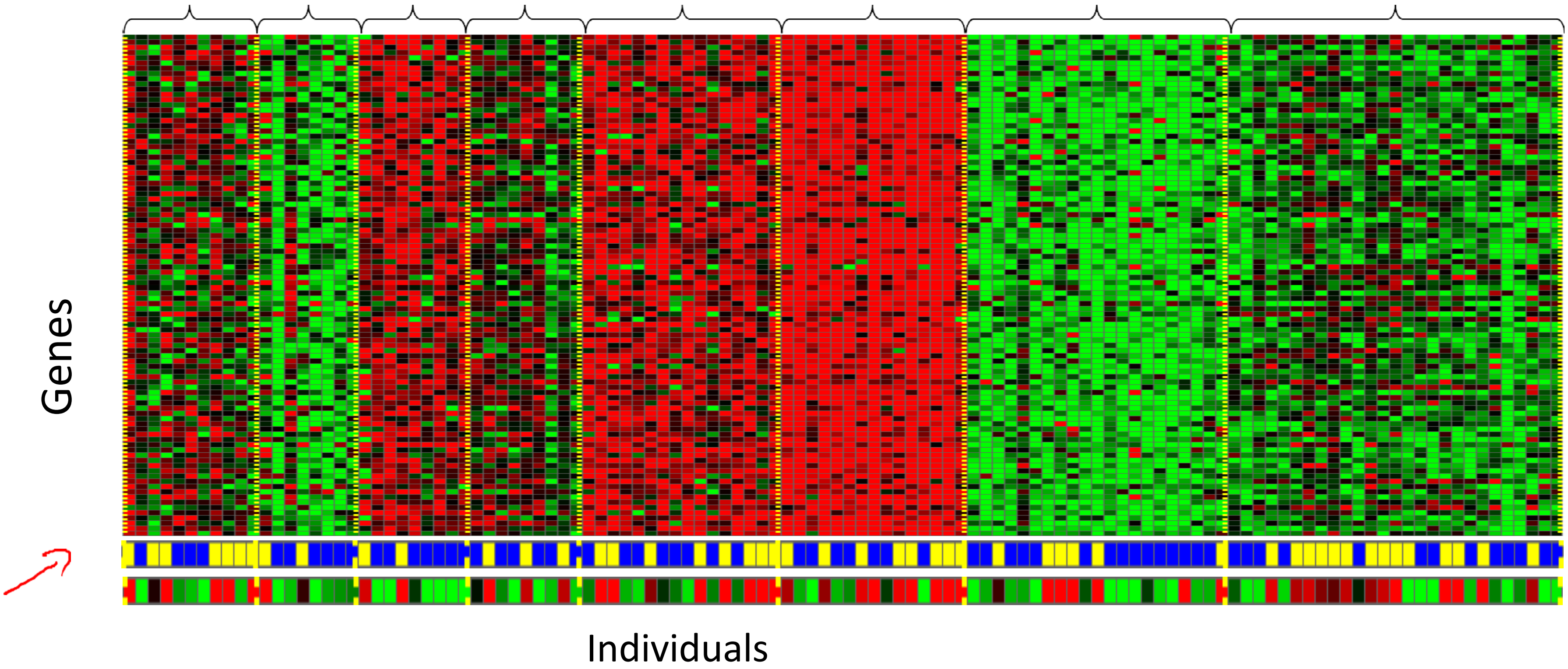
Spotlight

Sarkozy rages at EU 'humiliation'
Financial Times - Peggy Hollinger - Sep 16, 2010


Google News
news.google.com
Top Stories
White House official denies Tea Party-focused ad campaign
US Stocks Climb After Recession Called Over, Homebuilders Gain
BP Oil Well, Site of National Catastrophe, Dies at One
Vanity Fair - Juli Weiner - 22 minutes ago
The BP oil well, site of the Deepwater Horizon explosion that led to the worst oil spill in US history, died today at one year old.
all 2,292 news articles »
blogs.wsj.com/source/2010/09/20/bp-kills-macondo-but-its-legacy-lives-on/
THE WALL STREET JOURNAL.
THE SOURCE
SEPTEMBER 20, 2010, 12:44 PM GMT
BP Kills Macondo, But Its Legacy Lives On
By James Herron
BP confirmed late Sunday that the Macondo well that leaked almost five million barrels of oil into the Gulf of Mexico has been permanently sealed, but the well will continue to affect BP and the wider oil industry for many years.
The most immediate worry for BP and its shareholders is how the authorities will apportion blame for the spill.
Fire boat response crews battled the blazing remnants of the off shore oil
edition.cnn.com/2010/US/09/20/gulf.oil.disaster/
CNN
Allen: Well is dead, but much Gulf Coast work remains
By the CNN Wire Staff
September 20, 2010 -- Updated 1317 GMT (2117 HKT)
Click to play
What next for Gulf oil spill?
STORY HIGHLIGHTS
(CNN) -- The ruptured Macondo well, a mile under the Gulf of
www.guardian.co.uk/environment/2010/sep/20/bp-oil-spill-dee
guardian.co.uk
BP oil spill cost hits nearly $10bn
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
Julia Kollewe
guardian.co.uk, Monday 20 September 2010 08.33 BST
BP's costs for the Deepwater Horizon disaster have hit $10bn. Photograph: Ho/Reuters

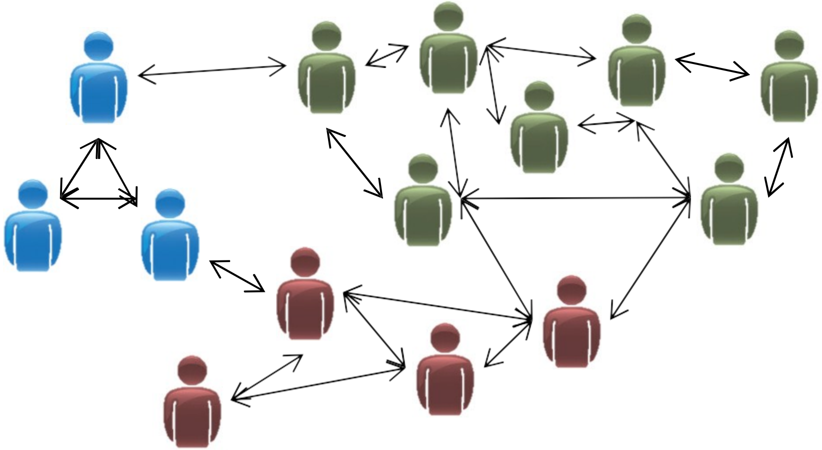


[Source: Daphne Koller]

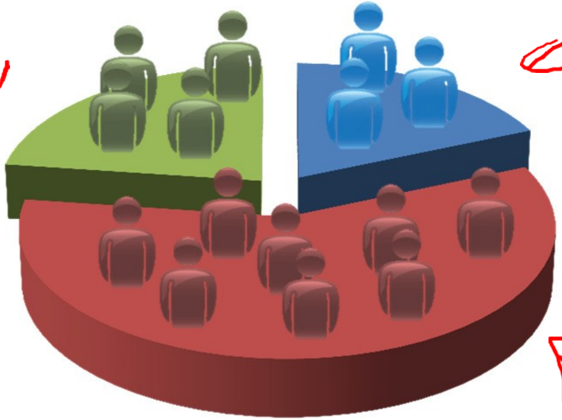


[Source: Daphne Koller]

Organize computing clusters



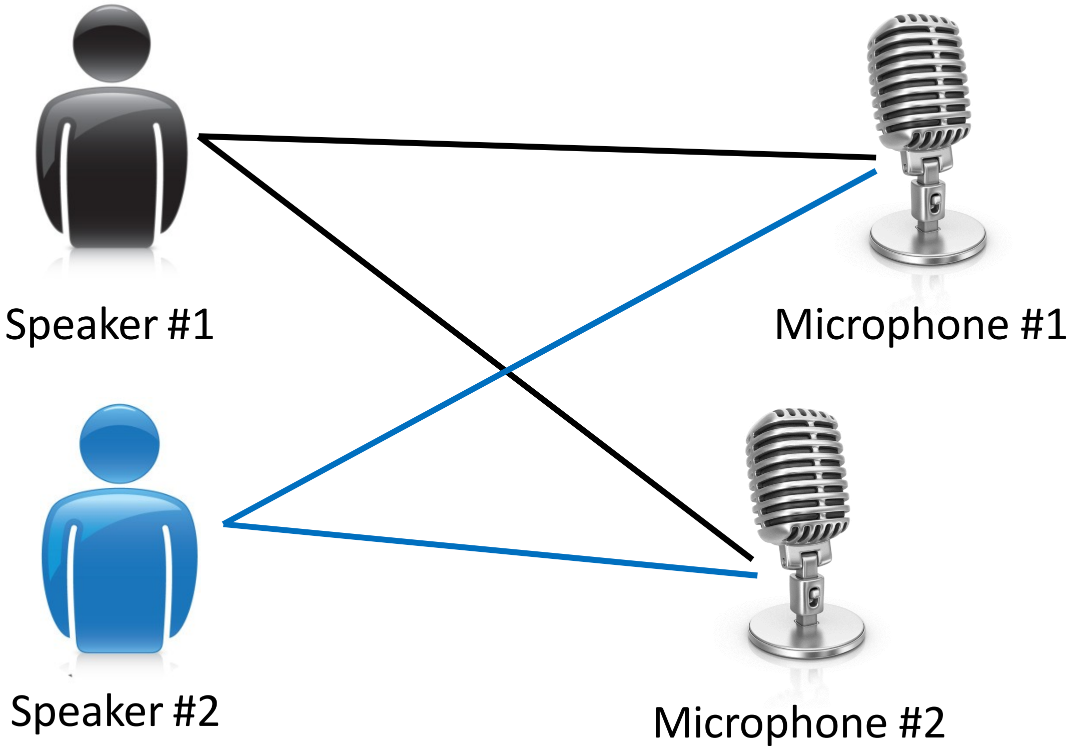
Social network analysis



Market segmentation

Astronomical data analysis

Cocktail party problem



Cocktail party problem algorithm

```
[W,s,v] = svd((repmat(sum(x.*x,1),size(x,1),1).*x)*x');
```

[Source: Sam Roweis, Yair Weiss & Eero Simoncelli]

Of the following examples, which would you address using an <u>unsupervised</u> learning algorithm? (Check all that apply.)

- ☐ Given email labeled as spam/not spam, learn a spam filter.
- ☐ Given a set of news articles found on the web, group them into set of articles about the same story.
- ☐ Given a database of customer data, automatically discover market segments and group customers into different market segments.
- ☐ Given a dataset of patients diagnosed as either having diabetes or not, learn to classify new patients as having diabetes or not.

Of the following examples, which would you address using an <u>unsupervised</u> learning algorithm? (Check all that apply.)

- ☐ Given email labeled as spam/not spam, learn a spam filter.
- ☐ Given a set of news articles found on the web, group them into set of articles about the same story.
- ☐ Given a database of customer data, automatically discover market segments and group customers into different market segments.
- ☐ Given a dataset of patients diagnosed as either having diabetes or not, learn to classify new patients as having diabetes or not.

Quiz 1 – ML Basics

1. A computer program is said to learn from experience E with respect to some Task T and some performance measure P if its performance on T, as measured by P, improves with experience E. Supposed we feed a learning algorithm a lot of historical weather data, and have it learn to predict weather. What would be a reasonable choice for P?

- o The process of the algorithm examining a large amount of historical weather data.
- o The weather prediction task.
- o The probability of it correctly predicting a future date's weather.
- o None of these.

2. Suppose you are working on weather prediction, and use a learning algorithm to predict tomorrow's temperature (in degrees Centigrade/Fahrenheit). Would you treat this as a classification or a regression problem?

- o Regression
- o Classification

3. Suppose you are working on stock market prediction, and you would like to predict the price of a particular stock tomorrow (measured in dollars). You want to use a learning algorithm for this. Would you treat this as a classification or a regression problem?

- o Regression
- o Classification

4. Some of the problems below are best addressed using a supervised learning algorithm, and the others with an unsupervised learning algorithm. Which of the following would you apply supervised learning to? (Select all that apply.) In each case, assume some appropriate dataset is available for your algorithm to learn from.

- ❑ Take a collection of 1000 essays written on the US Economy, and find a way to automatically group these essays into a small number of groups of essays that are somehow "similar" or "related".
- ❑ Examine a large collection of emails that are known to be spam email, to discover if there are sub-types of spam mail.
- ❑ Given 50 articles written by male authors, and 50 articles written by female authors, learn to predict the gender of a new manuscript's author (when the identity of this author is unknown).
- ❑ Given historical data of children's ages and heights, predict children's height as a function of their age.

5. Which of these is a reasonable definition of machine learning?

- o Machine learning is the field of allowing robots to act intelligently.
- o Machine learning is the science of programming computers.
- o Machine learning learns from labeled data.
- o Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed.

Quiz 1 – ML Basics

1. A computer program is said to learn from experience E with respect to some Task T and some performance measure P if its performance on T, as measured by P, improves with experience E. Supposed we feed a learning algorithm a lot of historical weather data, and have it learn to predict weather. What would be a reasonable choice for P?

- The process of the algorithm examining a large amount of historical weather data.
- The weather prediction task.
- The probability of it correctly predicting a future date's weather.
- None of these.

2. Suppose you are working on weather prediction, and use a learning algorithm to predict tomorrow's temperature (in degrees Centigrade/Fahrenheit). Would you treat this as a classification or a regression problem?

- Regression
- Classification

3. Suppose you are working on stock market prediction, and you would like to predict the price of a particular stock tomorrow (measured in dollars). You want to use a learning algorithm for this. Would you treat this as a classification or a regression problem?

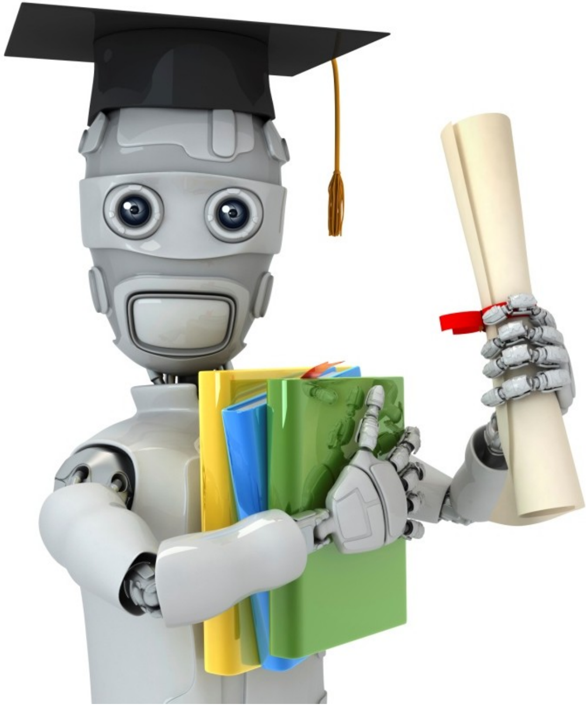
- Regression
- Classification

4. Some of the problems below are best addressed using a supervised learning algorithm, and the others with an unsupervised learning algorithm. Which of the following would you apply supervised learning to? (Select all that apply.) In each case, assume some appropriate dataset is available for your algorithm to learn from.

- [ ] Take a collection of 1000 essays written on the US Economy, and find a way to automatically group these essays into a small number of groups of essays that are somehow "similar" or "related".
- [ ] Examine a large collection of emails that are known to be spam email, to discover if there are sub-types of spam mail.
- [ ] Given 50 articles written by male authors, and 50 articles written by female authors, learn to predict the gender of a new manuscript's author (when the identity of this author is unknown).
- [ ] Given historical data of children's ages and heights, predict children's height as a function of their age.

5. Which of these is a reasonable definition of machine learning?

- Machine learning is the field of allowing robots to act intelligently.
- Machine learning is the science of programming computers.
- Machine learning learns from labeled data.
- Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed.

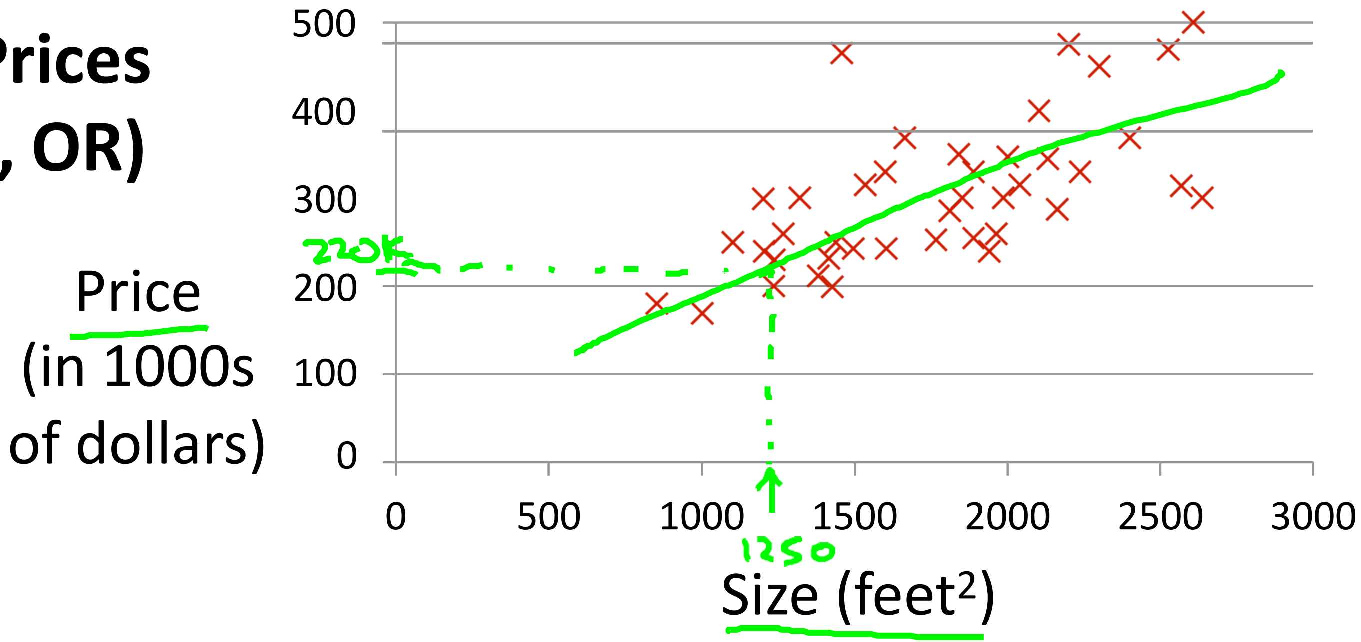


Machine Learning

Linear regression with one variable

Model representation

Housing Prices (Portland, OR)



<u>Supervised Learning</u>

Given the "right answer" for each example in the data.

<u>Regression Problem</u>

Predict real-valued output

Classification: Discrete-valued output

Training set of housing prices (Portland, OR)

Size in feet^2 (x)	Price ($) in 1000's (y)
2104	460
1416	232
1534	315
852	178
…	…

$m = 47$

Notation:

- **m** = Number of training examples
- **x**'s = "input" variable / features
- **y**'s = "output" variable / "target" variable

(x, y) - one training example

$(x^{(i)}, y^{(i)})$ - i^{th} training example

$x^{(1)} = 2104$

$x^{(2)} = 1416$

$y^{(1)} = 460$

Consider the training set shown below. $(x^{(i)}, y^{(i)})$ is the i^{th} training example. What is $y^{(3)}$?

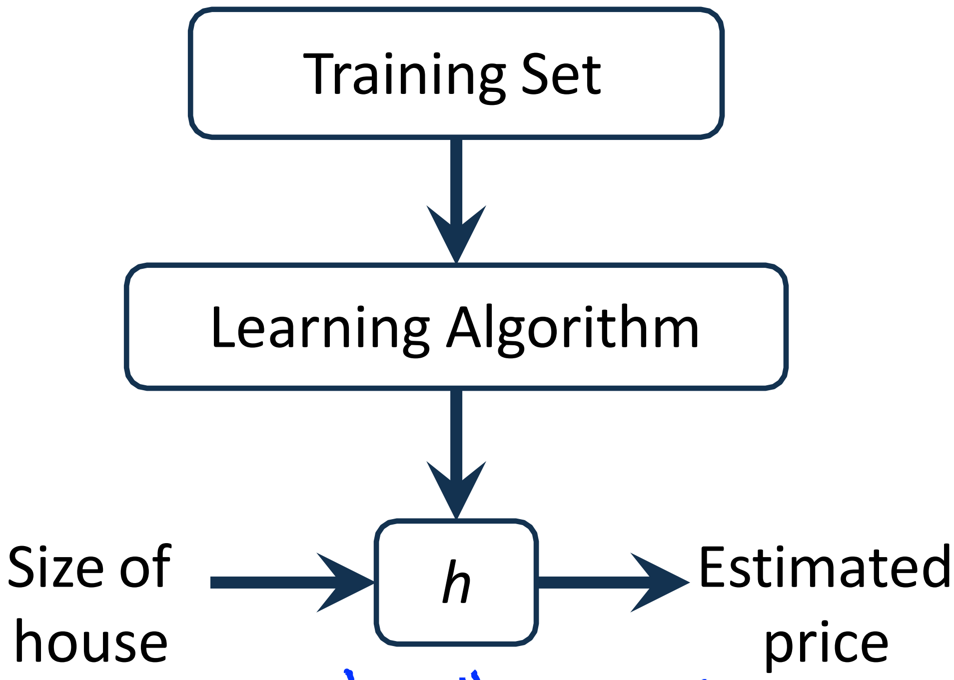
Size in feet^2(x)	Price($) in 1000's(y)
2104	460
1416	232
1534	315
852	178
….	….

- 1416
- 1534
- 315
- 0

Consider the training set shown below. $(x^{(i)}, y^{(i)})$ is the i^{th} training example. What is $y^{(3)}$?

Size in feet^2(x)	Price($) in 1000's(y)
2104	460
1416	232
1534	315
852	178
….	….

- ○ 1416
- ○ 1534
- ○ 315
- ○ 0



x

hypothesis

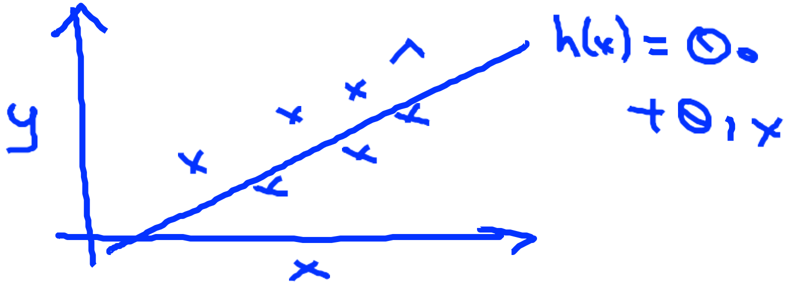
(estimated value of y)

h maps from x's to y's.

How do we represent *h* ?

$h_\theta(x) = \underline{\theta_0 + \theta_1 x}$

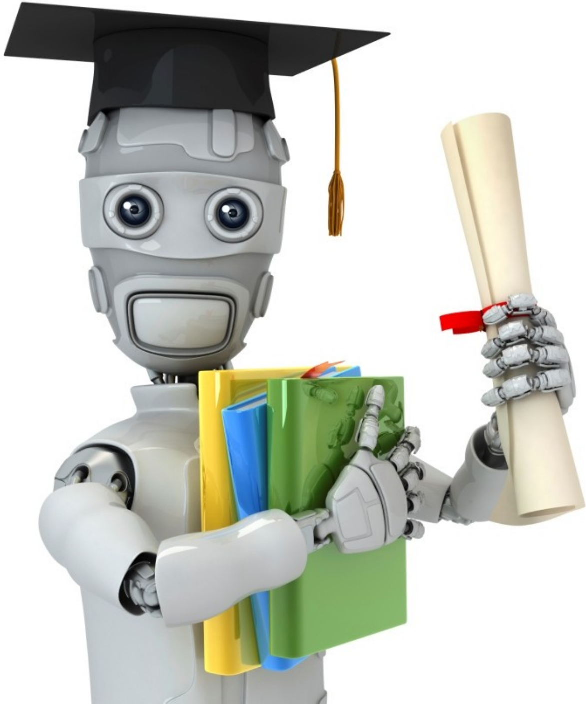
Shorthand: $h(x)$



Linear regression with one variable. (x)

Univariate linear regression.

└ one variable



Machine Learning

Linear regression with one variable

Cost function

Training Set

Size in feet^2 (x)	Price ($) in 1000's (y)
2104	460
1416	232
1534	315
852	178
…	…

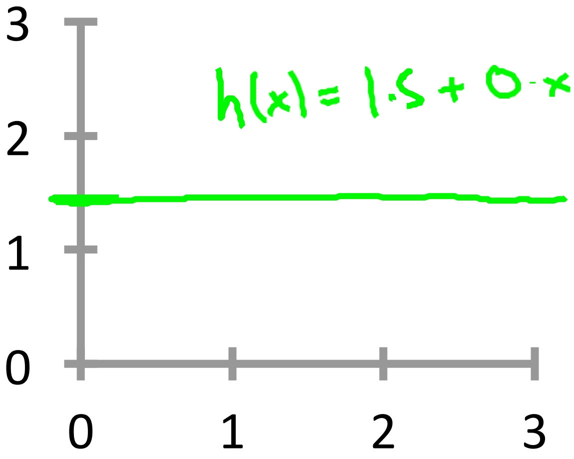
$m = 47$

Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$

θ_i's: Parameters

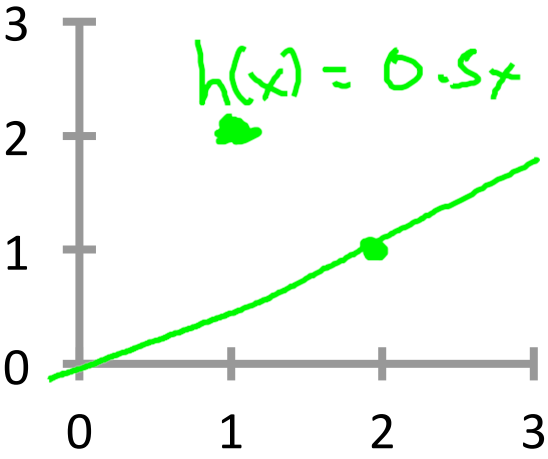
How to choose θ_i's ?

$$h_\theta(x) = \theta_0 + \theta_1 x$$



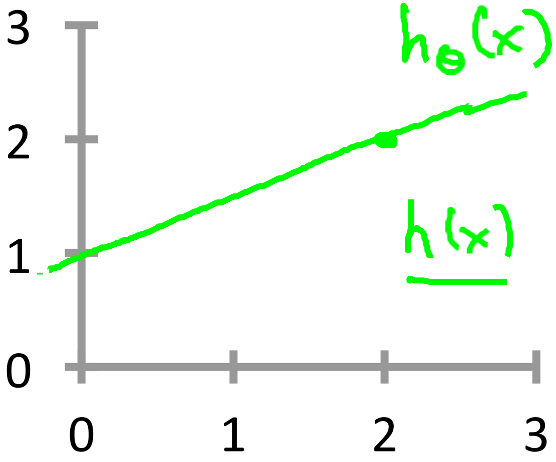
$\theta_0 = 1.5$

$\theta_1 = 0$



$\theta_0 = 0$

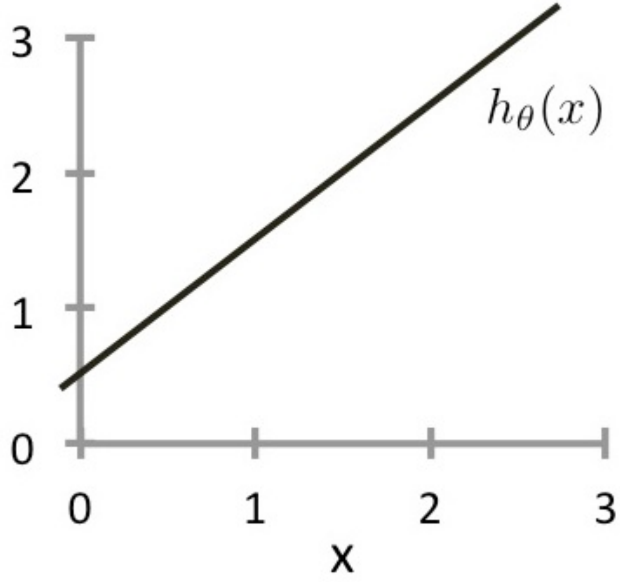
$\theta_1 = 0.5$



$\theta_0 = 1$

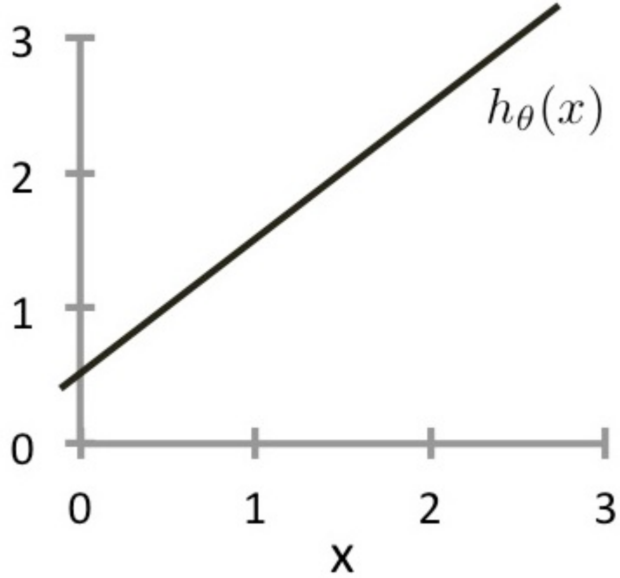
$\theta_1 = 0.5$

Consider the plot below of $h_\Theta(x) = \Theta_0 + \Theta_1 x$. What are Θ_0 and Θ_1?

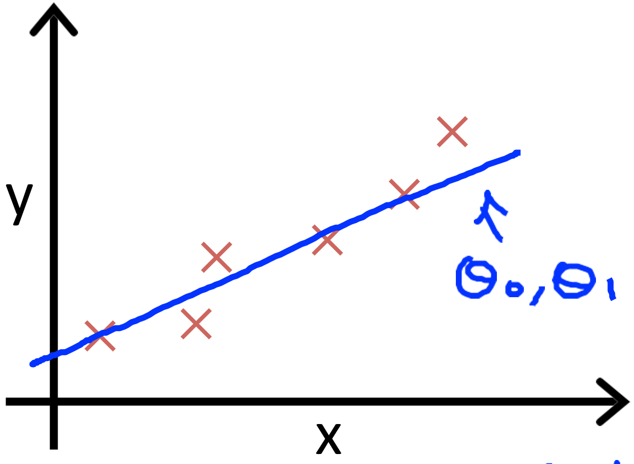


- $\Theta_0 = 0, \Theta_1 = 1$
- $\Theta_0 = 0.5, \Theta_1 = 1$
- $\Theta_0 = 1, \Theta_1 = 0.5$
- $\Theta_0 = 1, \Theta_1 = 1$

Consider the plot below of $h_\Theta(x) = \Theta_0 + \Theta_1 x$. What are Θ_0 and Θ_1?



- $\Theta_0 = 0, \Theta_1 = 1$
- $\Theta_0 = 0.5, \Theta_1 = 1$
- $\Theta_0 = 1, \Theta_1 = 0.5$
- $\Theta_0 = 1, \Theta_1 = 1$



$(x^{(i)}, y^{(i)})$

\# training examples

$$\underset{\theta_0,\theta_1}{\text{minimize}} \ \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

$$h_\theta(x^{(i)}) = \theta_0 + \theta_1 x^{(i)}$$

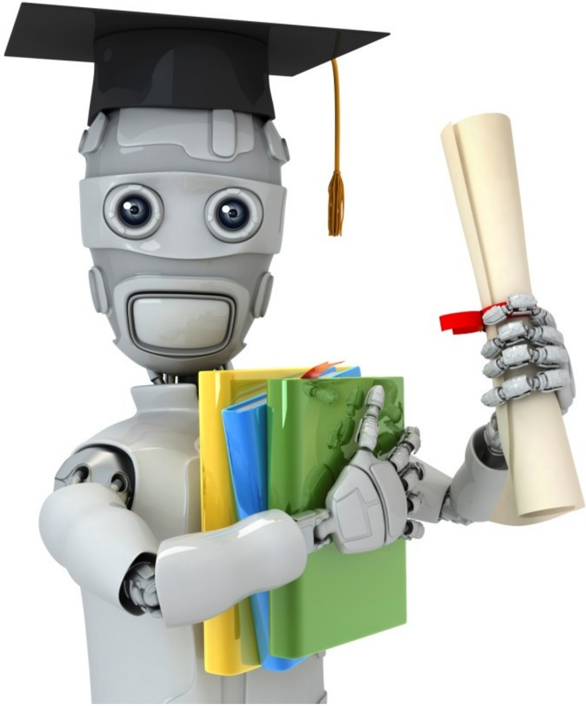
Idea: Choose θ_0, θ_1 so that $h_\theta(x)$ is close to y for our training examples (x, y)

x, y

$$J(\theta_0,\theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

$$\underset{\theta_0,\theta_1}{\text{minimize}} \ \underbrace{J(\theta_0,\theta_1)}_{\text{Cost function}}$$

Squared error function



Machine Learning

Linear regression with one variable

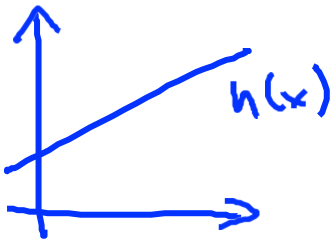
Cost function intuition I

Hypothesis:

$$h_\theta(x) = \theta_0 + \theta_1 x$$

Parameters:

$$\theta_0, \theta_1$$



Cost Function:

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2$$

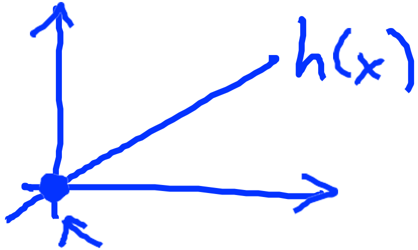
Goal: $\underset{\theta_0, \theta_1}{\text{minimize}} \; J(\theta_0, \theta_1)$

Simplified

$$h_\theta(x) = \theta_1 x$$

$\theta_0 = 0$

$$\theta_1$$



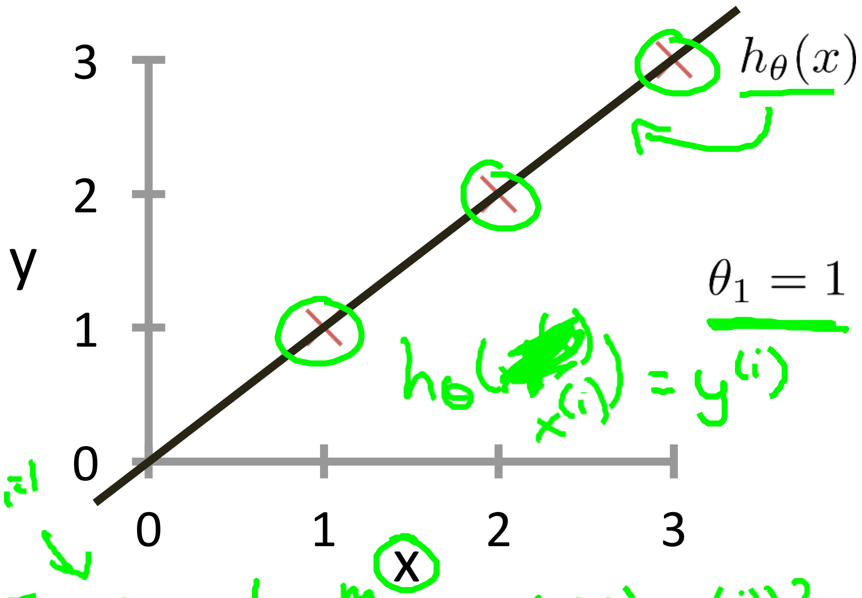
$$J(\theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2$$

$\theta_1, x^{(i)}$

$$\underset{\theta_1}{\text{minimize}} \; J(\theta_1)$$

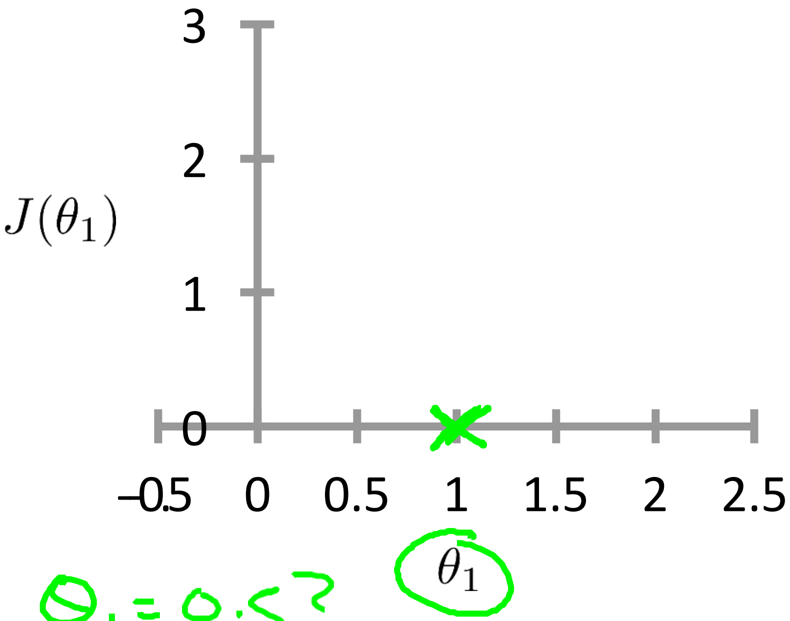
$\rightarrow h_\theta(x)$

(for fixed θ_1, this is a function of x)



$\rightarrow J(\theta_1)$

(function of the parameter θ_1)



$$J(\theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$$

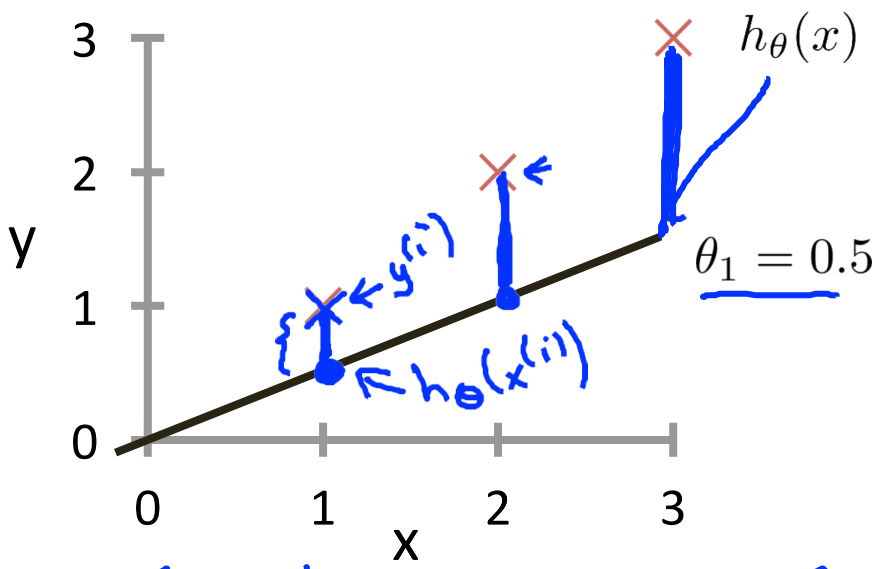
$$= \frac{1}{2m}\sum_{i=1}^{m}\left(\theta_1 x^{(i)} - y^{(i)}\right)^2 = \frac{1}{2m}\left(0^2+0^2+0^2\right) = 0^2$$

$\theta_1 = 0.5$?

$J(1) = 0$

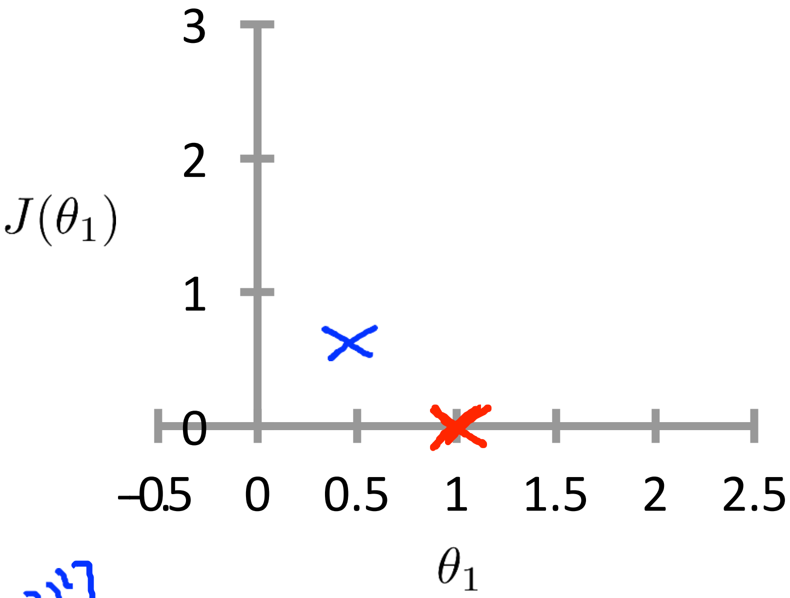
$h_\theta(x)$

(for fixed θ_1, this is a function of x)



$J(\theta_1)$

(function of the parameter θ_1)



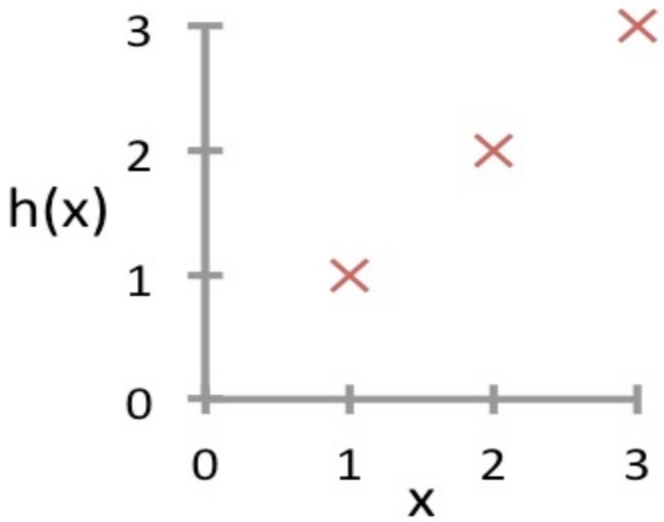
$J(0.5) = \frac{1}{2m}\left[(0.5-1)^2 + (1-2)^2 + (1.5-3)^2\right]$

$= \frac{1}{2\times 3}(3.5) = \frac{3.5}{6} \approx 0.58$

$\theta_1 = 0?$

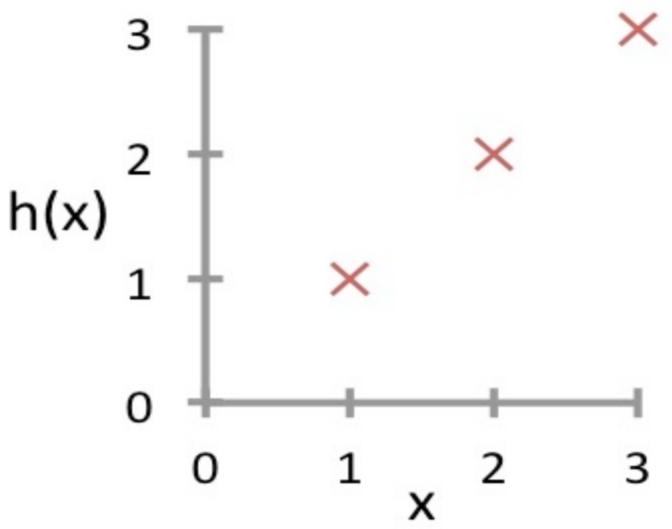
$J(0) = ?$

Suppose we have a training set with m=3 examples, plotted below. Our hypothesis representation is $h_\Theta(x) = \Theta_1 x$, with parameter Θ_1. The cost function $J(\Theta_1)$ is $J(\Theta_1) = (1/2m) \sum_{i=1}^{m} (h_\Theta(x^{(i)}) - y^{(i)})^2$. What is J(0)?



- 0
- 1/6
- 1
- 14/6

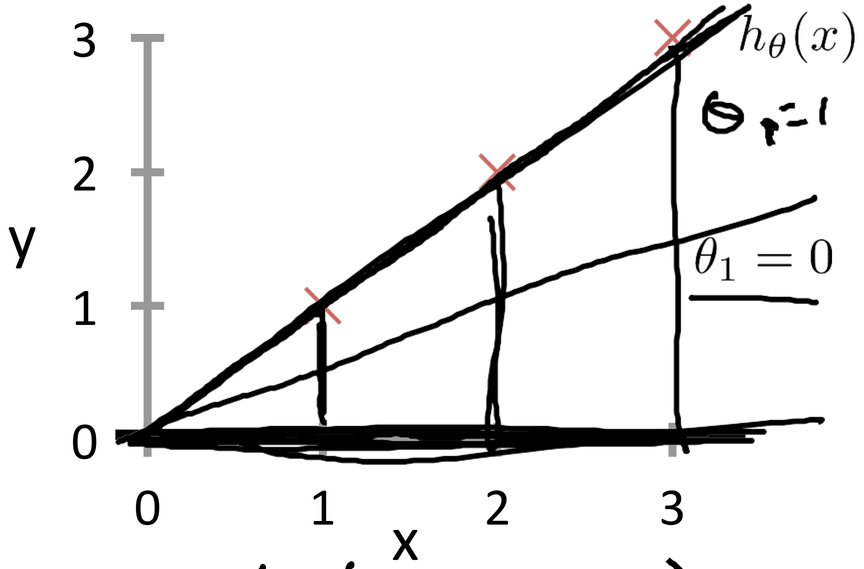
Suppose we have a training set with m=3 examples, plotted below. Our hypothesis representation is $h_\Theta(x) = \Theta_1 x$, with parameter Θ_1. The cost function $J(\Theta_1)$ is $J(\Theta_1) = (1/2m) \sum_{i=1}^{m} (h_\Theta(x^{(i)}) - y^{(i)})^2$. What is J(0)?



- 0
- 1/6
- 1
- 14/6

$h_\theta(x)$

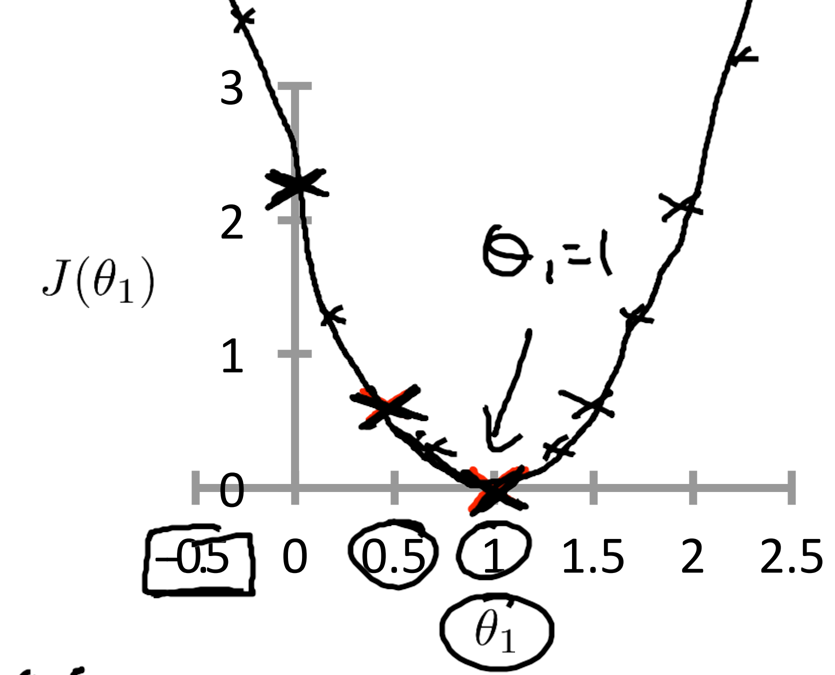
(for fixed θ_1, this is a function of x)



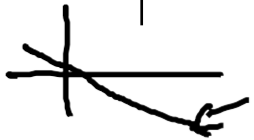
$$J(0) = \frac{1}{2m}\left(1^2 + 2^2 + 3^2\right)$$
$$= \frac{1}{6} \cdot 14 \approx 2.3$$

$J(\theta_1)$

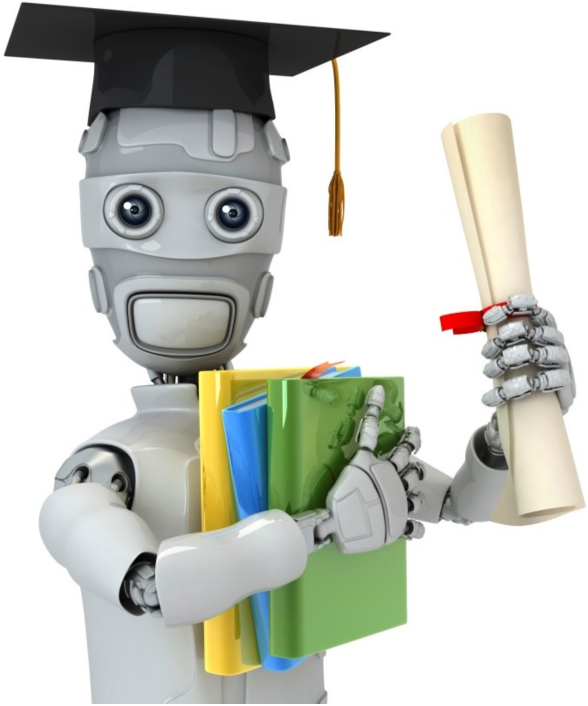
(function of the parameter θ_1)



$h(x) = -0.5x$



$$\boxed{\underset{\theta_1}{\text{minimize}}\ J(\theta_1)}$$



Machine Learning

Linear regression with one variable

Cost function intuition II

Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$

Parameters: $\underline{\theta_0, \theta_1}$

Cost Function: $J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)}\right)^2$

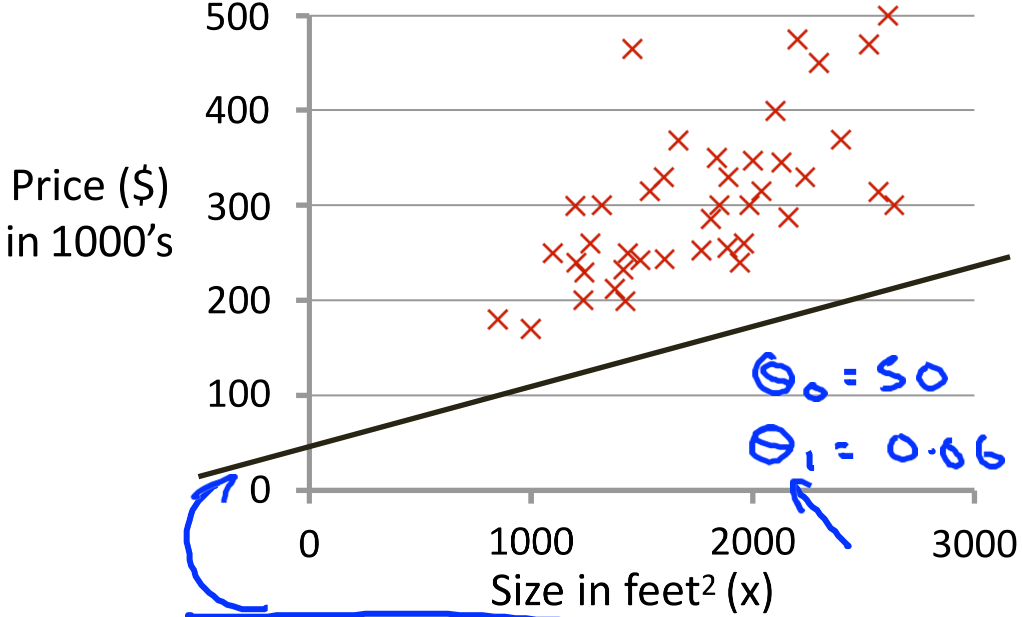
Goal: $\underset{\theta_0, \theta_1}{\text{minimize}}\, J(\theta_0, \theta_1)$

$h_\theta(x)$

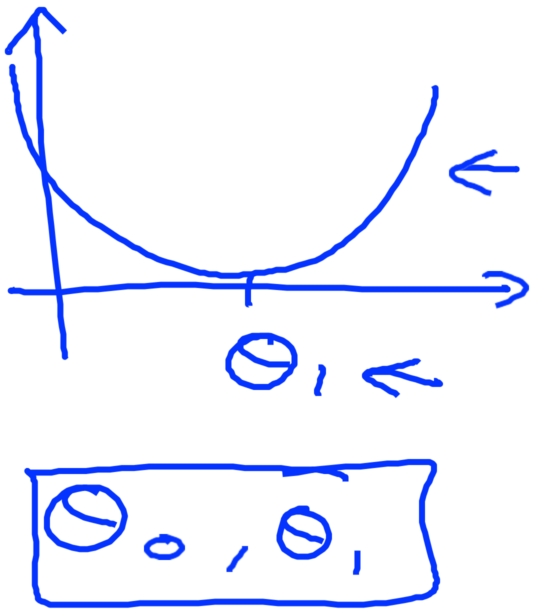
(for fixed θ_0, θ_1, this is a function of x)

$J(\theta_0, \theta_1)$

(function of the parameters θ_0, θ_1)

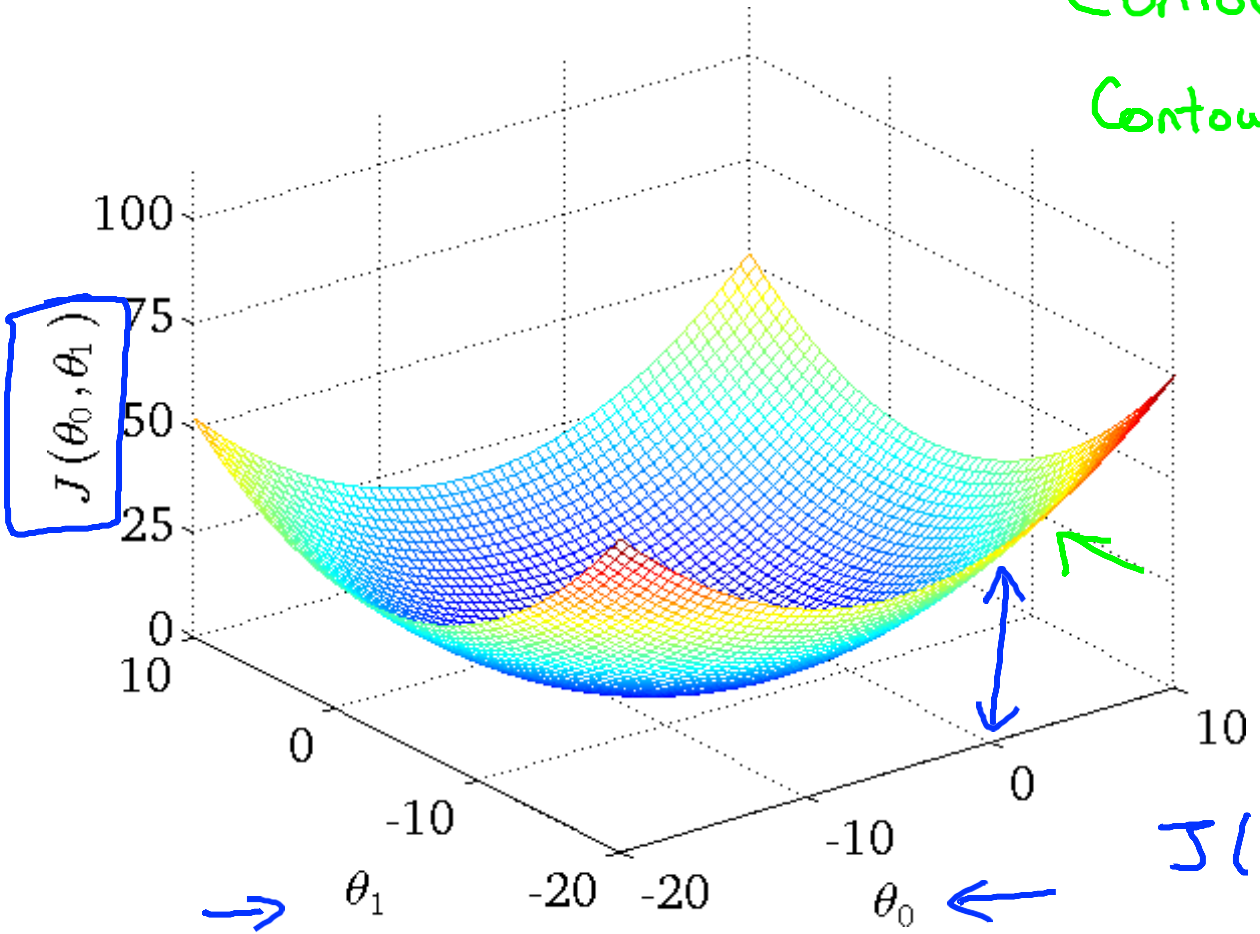


$$\boxed{h_\theta(x) = 50 + 0.06x}$$



Contour plots

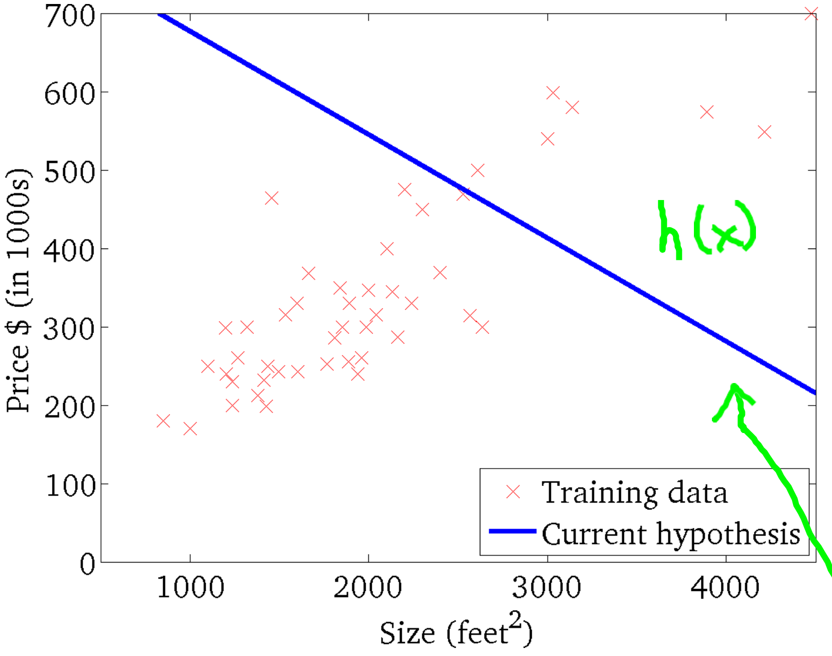
Contour figures.



$J(\theta_0, \theta_1)$

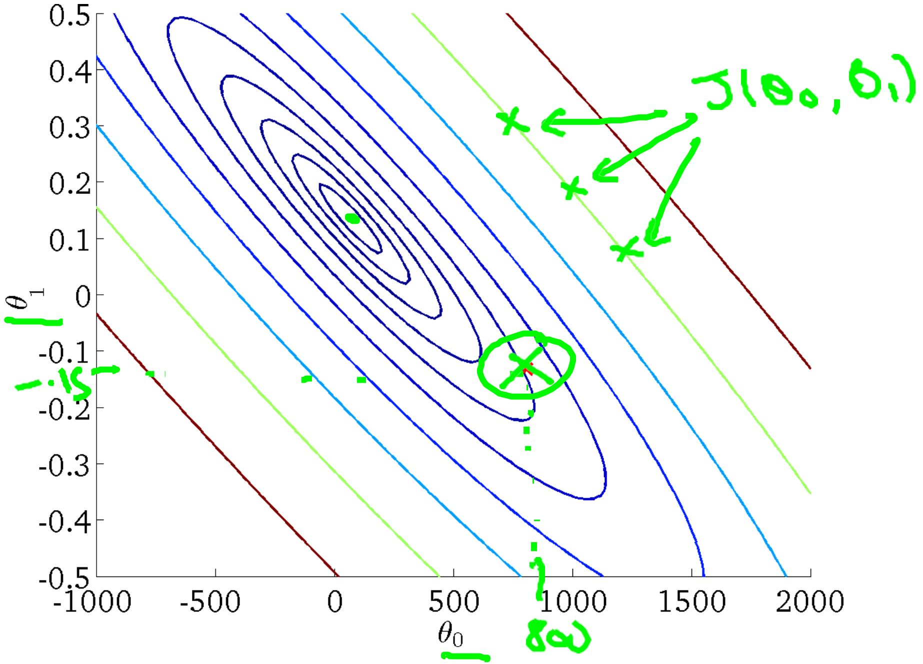
$h_\theta(x)$

(for fixed θ_0, θ_1, this is a function of x)



$J(\theta_0, \theta_1)$

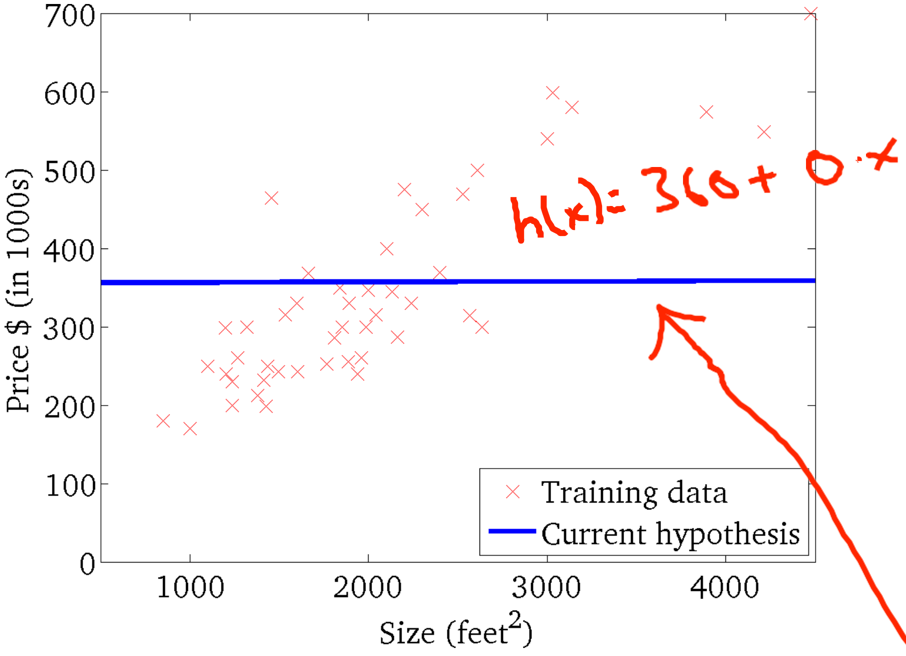
(function of the parameters θ_0, θ_1)



θ_0, θ_1

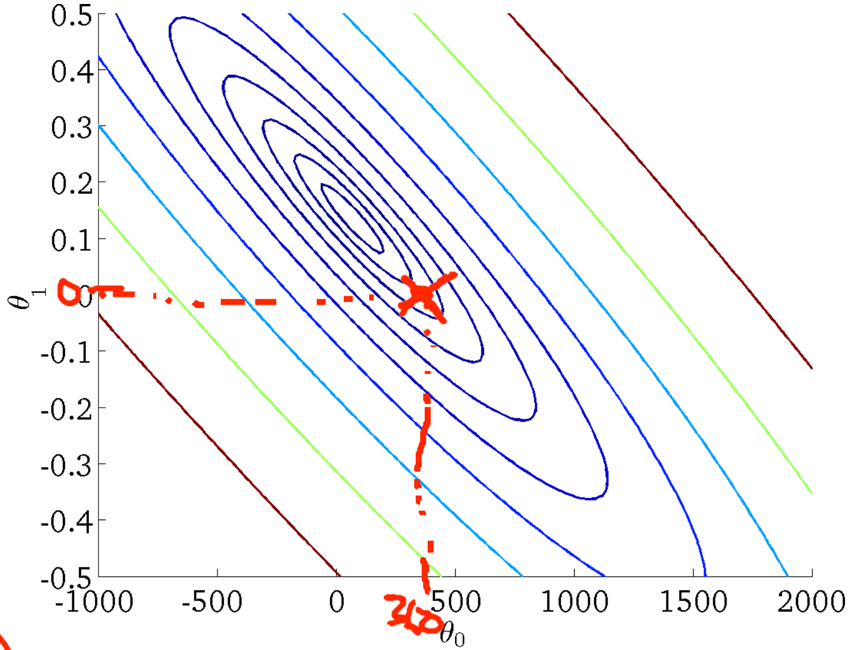
$$h_\theta(x)$$

(for fixed θ_0, θ_1, this is a function of x)



$$J(\theta_0, \theta_1)$$

(function of the parameters θ_0, θ_1)

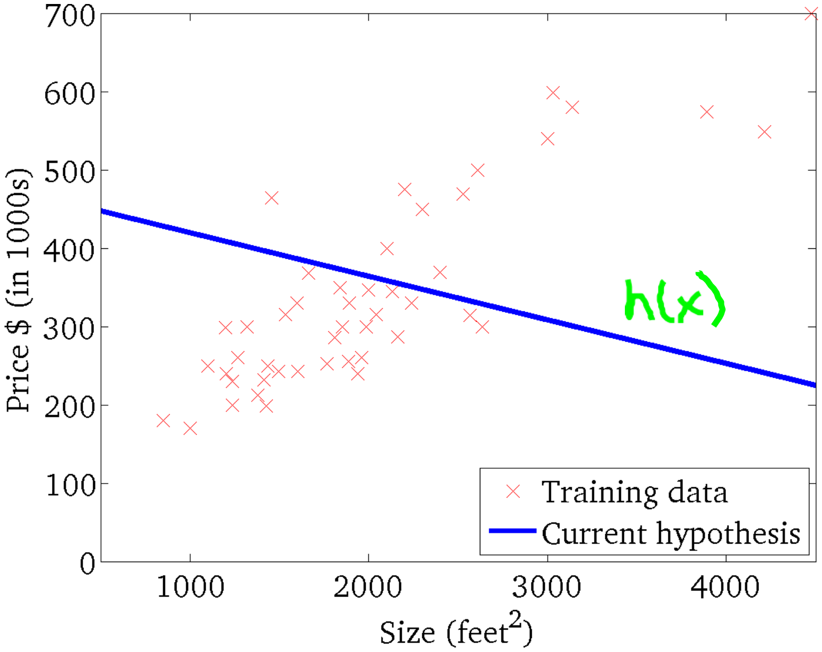


$\theta_0 = 360$

$\theta_1 = 0$

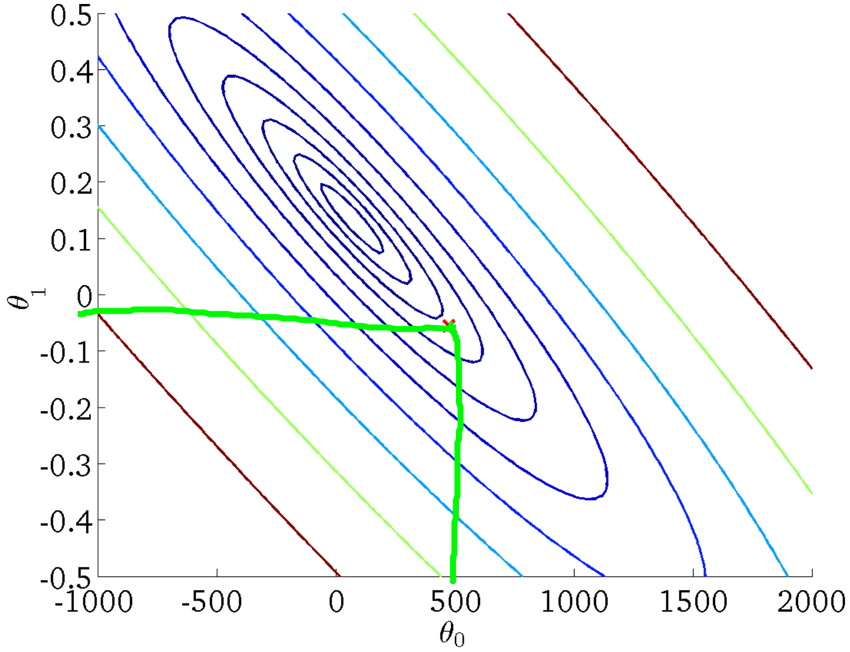
$h_\theta(x)$

(for fixed θ_0, θ_1, this is a function of x)



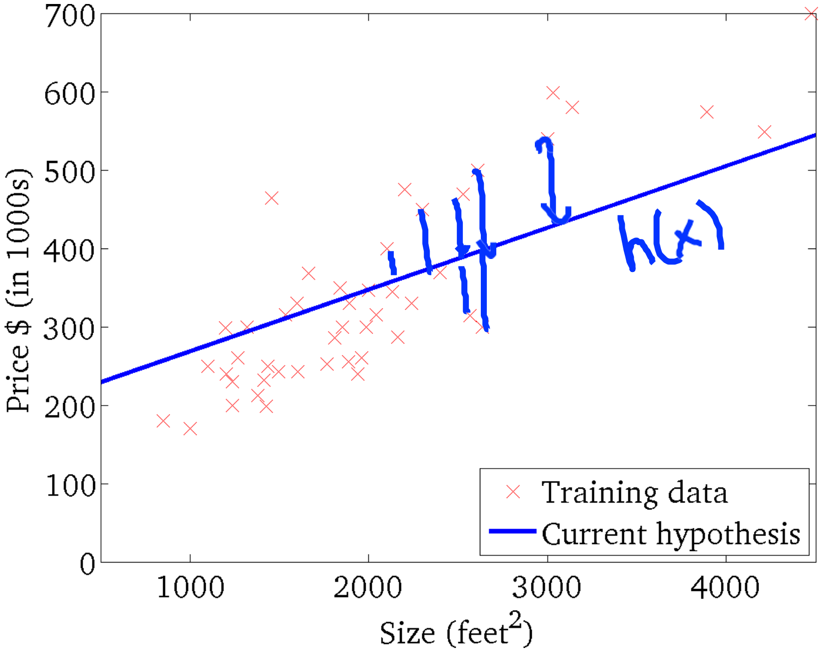
$J(\theta_0, \theta_1)$

(function of the parameters θ_0, θ_1)



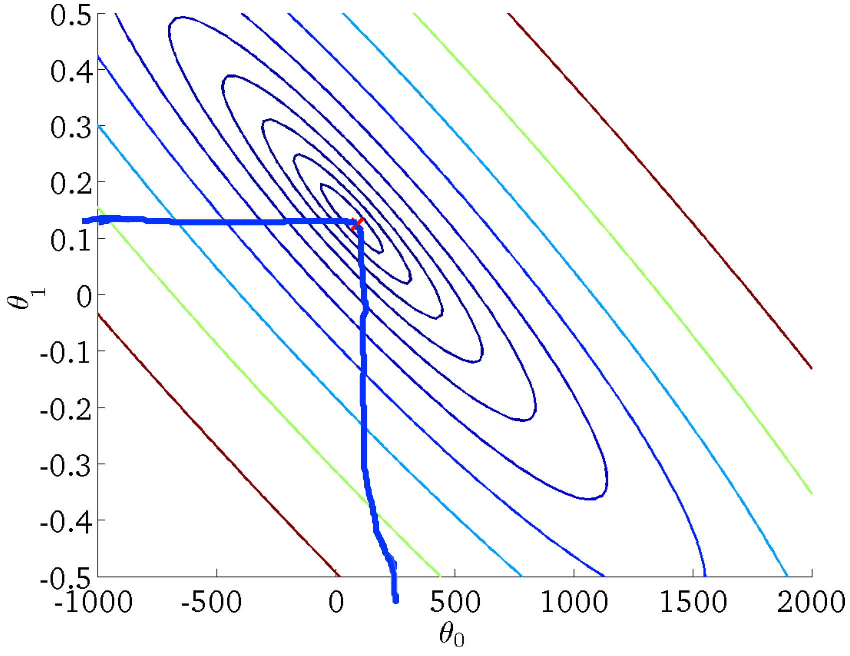
$h_\theta(x)$

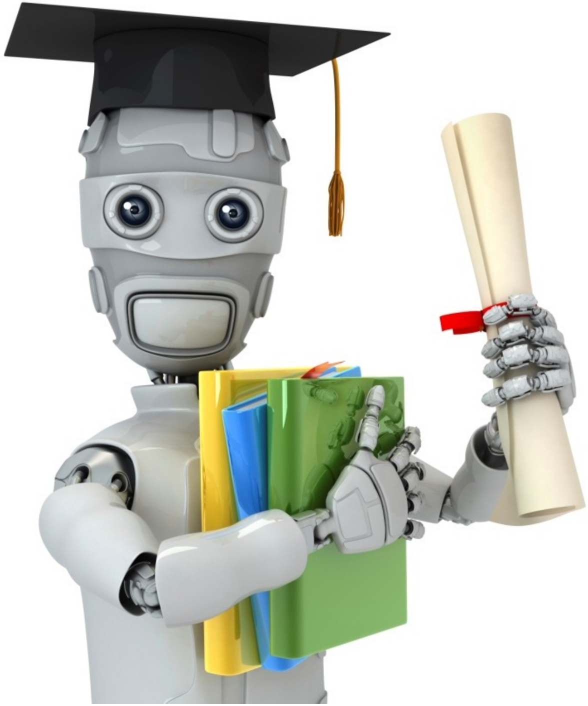
(for fixed θ_0, θ_1, this is a function of x)



$J(\theta_0, \theta_1)$

(function of the parameters θ_0, θ_1)





Machine Learning

Linear regression with one variable

Next Seminar
Gradient descent