

Machine Learning

Tran Thi Minh Hanh

hanhtran@dut.udn.vn

<https://sites.google.com/view/hanh-tran/lectures/machine-learning>

Goals

1. To provide students with a broad introduction to machine learning, data mining, and statistical pattern recognition.
2. To equip students with programming, teamwork, problem solving skills through project work.

References

- Christopher M. bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- *Machine Learning course*, Andrew Ng, Stanford University.

Grading

- Exercise: 20%
- Midterm Exam.: 20%
- Final Exam.: 60%

Contents

- **Introduction**
- **Supervised learning**
 - Linear regression
 - Classification
 - Neural network
 - Support Vector Machine
- **Unsupervised learning**
 - Clustering
 - PCA
- **Deep learning**

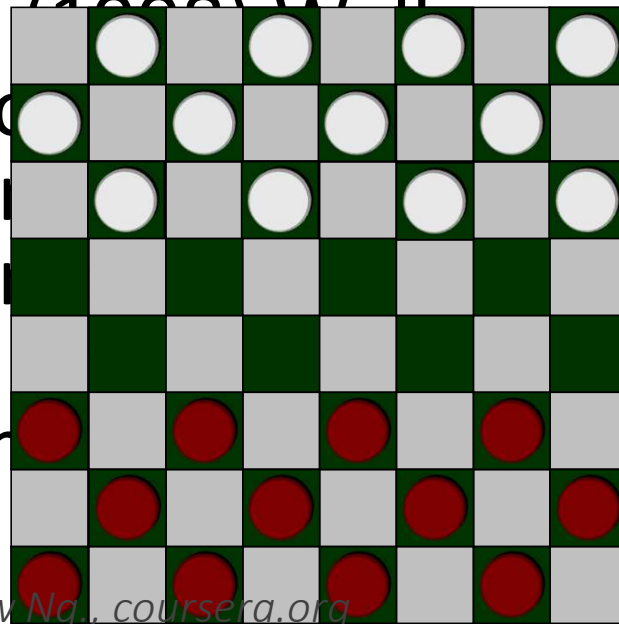
Introduction

- What is machine learning
- Supervised learning
- Unsupervised learning

Source: Machine Learning, Andrew Ng., coursera.org

Machine Learning definition

- Arthur Samuel (1959). Machine Learning: Field of study that gives computers the ability to learn without being explicitly programmed.
- Tom Mitchell (1997). What is 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 P on task T as measured by P , improves with experience.



Source: Machine Learning, Andrew Ng, coursera.org

“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.

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., 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 algorithms:

- Supervised learning

- Learn to predict an output when given an input vector.

- Unsupervised learning

- Discover a good internal representation of the input.

Others: Reinforcement learning, recommender systems.

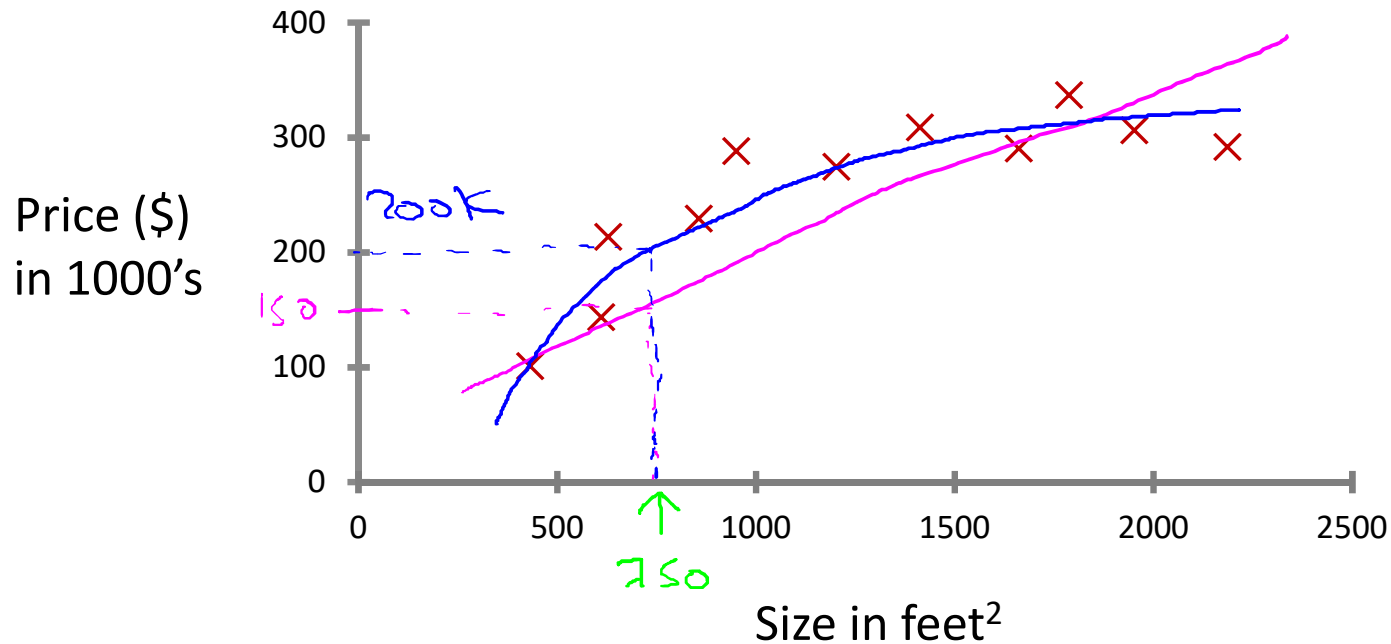
Introduction

- What is machine learning
- Supervised learning
- Unsupervised learning

Two types of supervised learning

- Each training data consists of an input vector x and a target output y .
- Regression: The target output is a real number or a vector of real numbers.
 - The price of a stock in 6 months time.
 - The temperature at noon tomorrow.
- Classification: The target output is a class label.

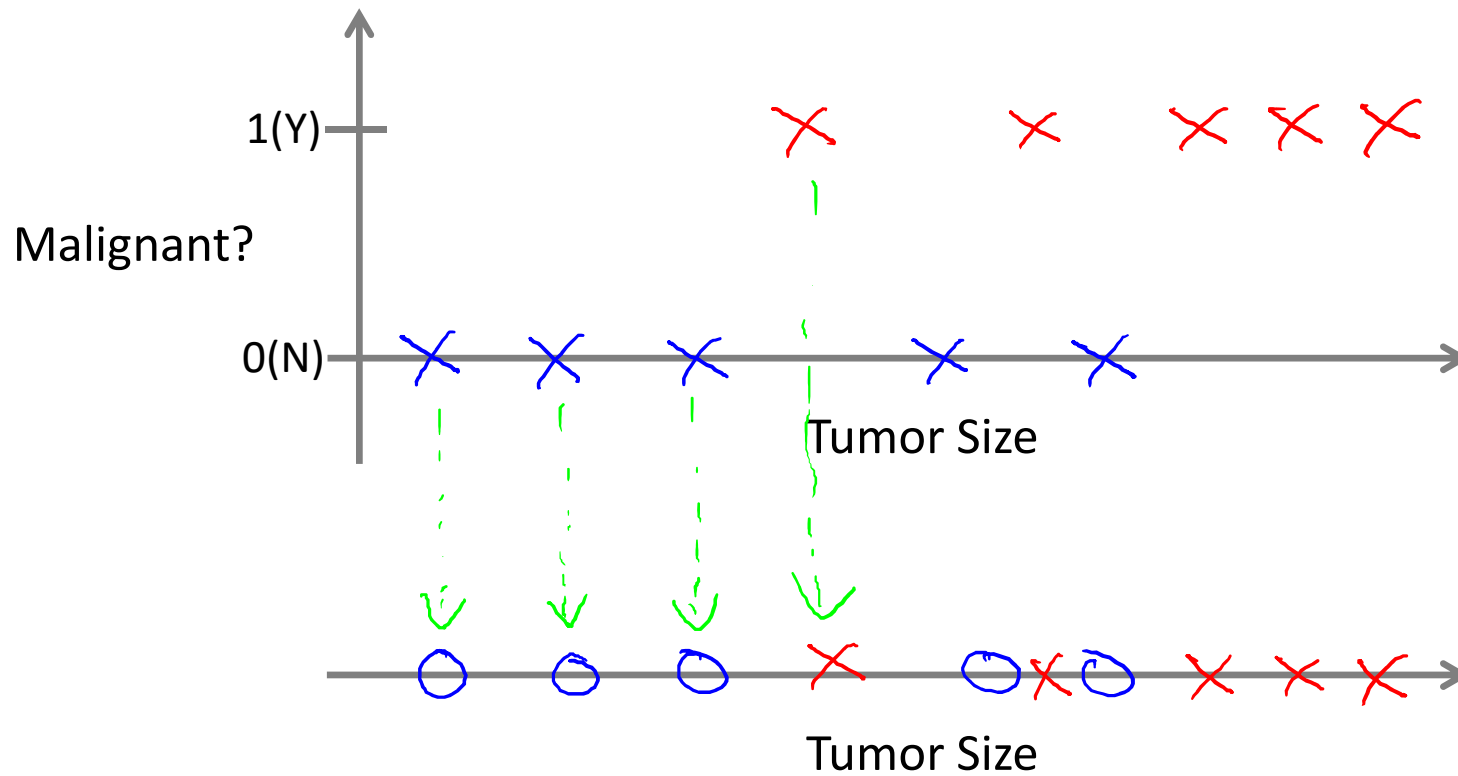
Housing price prediction.



Supervised Learning

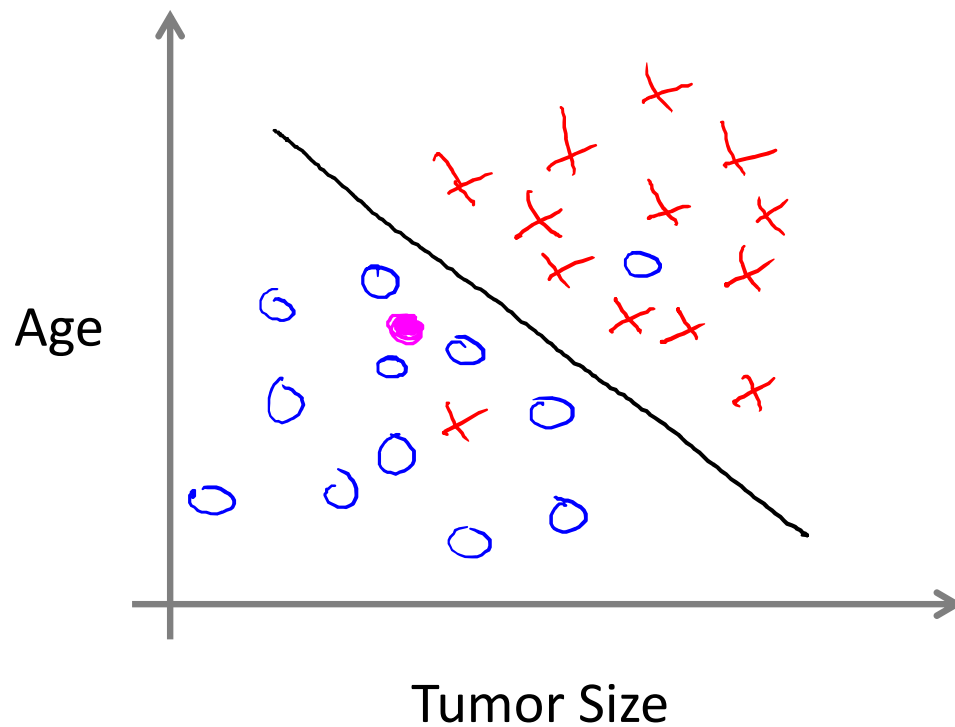
“right answers” given

Regression: Predict continuous
valued output (price)



Classification

Discrete valued
output (0 or 1)



- 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.

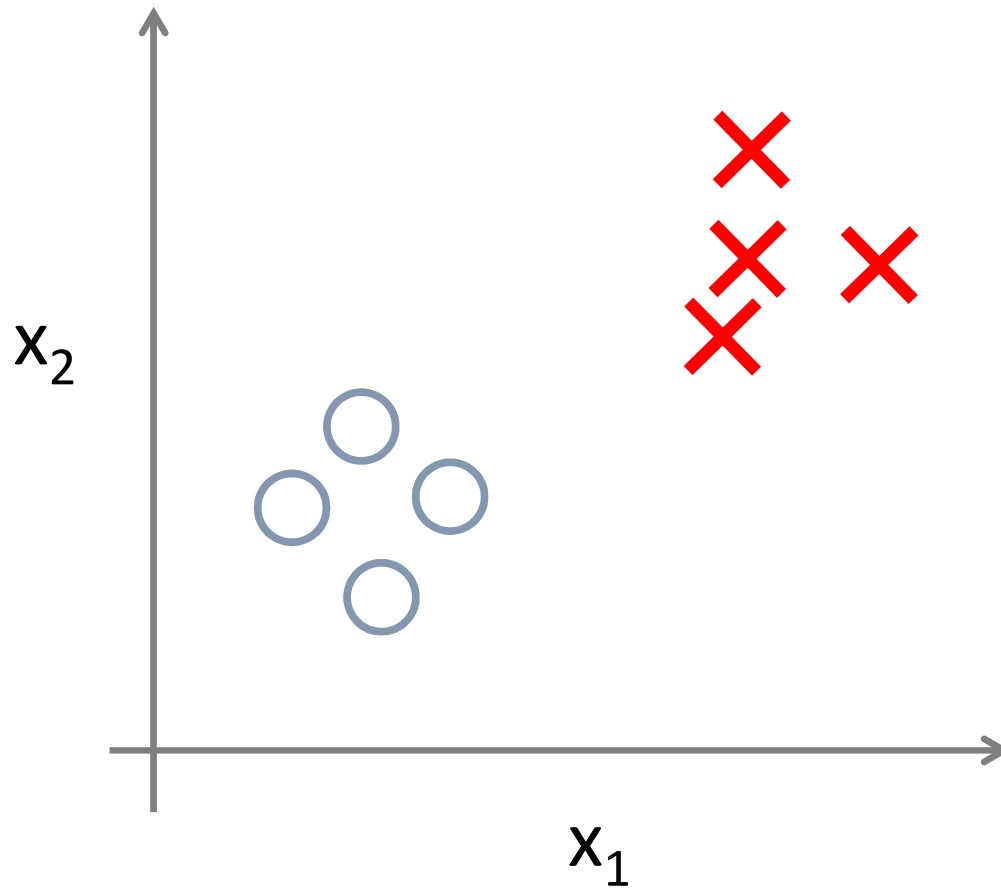
Treat problem 1 as a regression problem, problem 2 as a classification problem.

Treat both as regression problems.

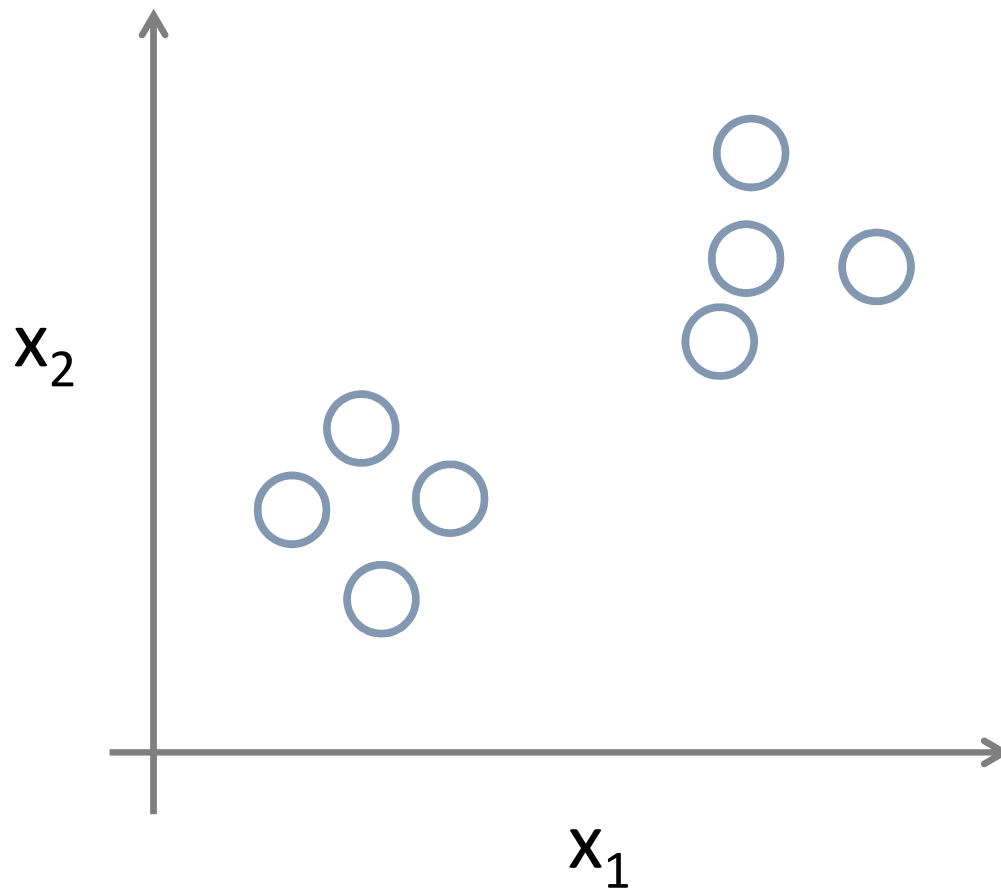
Introduction

- What is machine learning
- Supervised learning
- 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 »

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 »

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 »

CNN Interna...
MyFox Phila...
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

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
SciTech
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 lying 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

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 24 news articles

BP Oil Well, Site of National Catastrophe, Dies at One
Vanity Fair - Juli Wiener - 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

Recent

Recession officially ended in June 2009
CNNMoney - Chris Isidore - 39 minutes ago
Hurricane Igor lashes Bermuda
USA Today - Gerry Browne - 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 headhunter 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 20, 2010

BP Kills Macondo, But Its Legacy Lives On

THE WALL STREET JOURNAL

Log In • Register For Free • Subscribe Now, Get 2 Weeks Free

PREVIOUS NEXT

Financial Services Transport Leisure Insurance Oil & Gas Sport Caught on the Web Betting Technology

SEPTEMBER 20, 2010, 12:44 PM GMT

BP Kills Macondo, But Its Legacy Lives On

Search The Source SEARCH

Article Comments (2)

THE SOURCE HOME PAGE

Email Print Permalink Like 2 + More Text

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. BP's own investigation spread responsibility across

Fire boat response crews battled the blazing remnants of the off shore oil platform Macondo on April 21, 2010.

Associated Press

About The Source

Follow Us

The Source is WSJ.com Europe's home for rapid-fire analysis of the day's big business and finance stories. It is edited by Lauren Mills, based in London.

Most Recent

Articles Comments

- 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
- We Already Need a Season to Rascal III

Allen: Well is dead, but much Gulf Coast work remains

edition.cnn.com/2010/US/09/20/gulf.oil.disaster/

EDITION: INTERNATIONAL | U.S. | MEXICO | ARABIC

Set edition preference

Home Video World U.S. Africa Asia Europe Latin America Middle East Business W

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 Mexico off the Louisiana coast, has been pronounced dead

BP oil spill cost hits nearly \$10bn

www.guardian.co.uk/environment/2010/sep/20/bp-oil-spill-dee

Mobile site Sign in Register A Text larger smaller About us

guardian.co.uk

News Sport Comment Culture Business Money Life & style

Business BP

BP oil spill cost hits nearly \$10bn

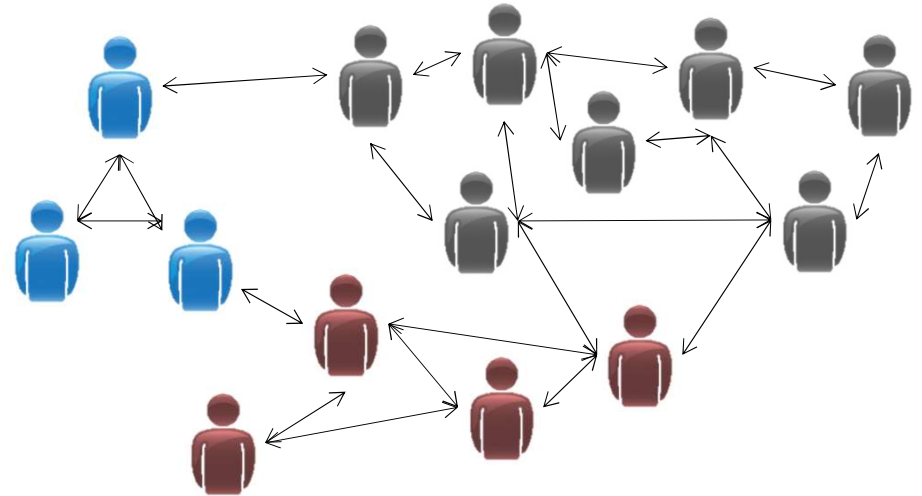
BP has set up a \$20bn compensation fund after the Deepwater Horizon disaster, which has so far paid out 19,000 claims totalling more than \$240m

Julia Kollewe
guardian.co.uk, Monday 20 September 2010 08.33 BST
Article history

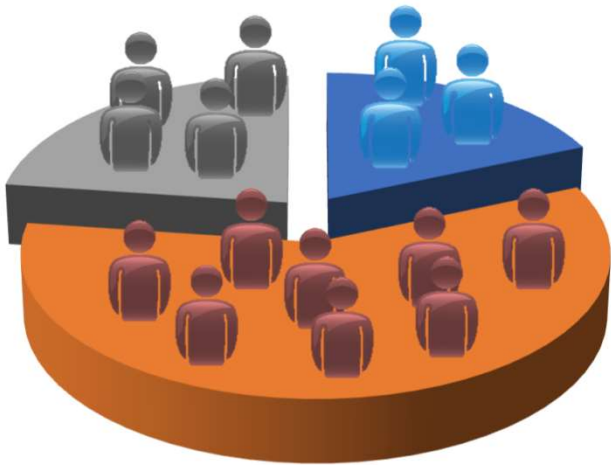
BP's costs for the Deepwater Horizon disaster have hit \$10bn. Photograph: Ho/Reuters



Organize computing clusters



Social network analysis



Market segmentation



Astronomical data analysis

Of the following examples, which would you address using an unsupervised 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.