

Machine learning

Presented by : Dr. Hanaa Bayomi



Lecture 1 : Introduction

Prof. Abeer ElKorany

Faculty Of Computers and Artificial intelligence

Cairo University

a.korani@fci-cu.edu.eg

Course Overview

- **Course Description**
- **Assessment**

Guidelines and Rules

- Keep your Mobile phones silent please.
- Alert me if I went faster than you could capture the content.
- Additional bonus for impressive answers

Course Overview

- Introduction to machine learning
- Linear regression
- Logistic regression
- Naive Bayes
- KNN
- Decision Tree
- Support vector machines
- Neural Networks:
- Clustering

REFERENCES

- T. Mitchell, *Machine Learning*, McGraw-Hill
- Peter Flach, **Machine Learning. The Art and Science of Algorithms that Make Sense of Data.**
- John D. Kelleher, *Fundamental of Machine Learning for predictive Data Analytic.*

Grading

Type	grades
▪ 3 Assignments (regression, SVM, DT)	16
▪ Mid term Exam	12
▪ project	12
▪ Final exam	60
Total	100

What is learning?

- "The activity or process of gaining knowledge or skill by studying, practicing, being taught, or experiencing something." Merriam Webster dictionary
- "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E ." Tom Mitchell
- "Learning is any process by which a system improves performance from experience." Herbert Simon

➤ In other words

- instead of the programmer writing explicit rules for how to solve a given problem, the programmer instructs the computer how to learn from examples
- in many cases the computer program can even become better at the task than the programmer is!

What is machine learning?

➤ Definition:

Machine = computer, computer program (in this course)

Learning = improving performance on a given task, based on experience / examples

Tom Mitchell: Algorithms that:

- improve their performance P
- at task T
- with experience E

A well-defined machine learning task is given by (P, E, T)

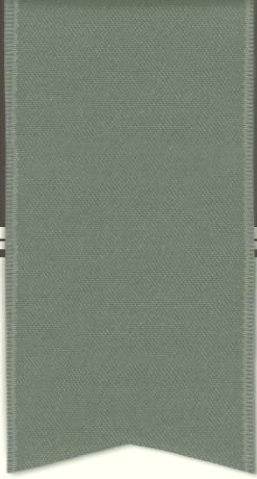
What is machine learning?

Traditional Programming



Machine Learning





EXAMPLES

Example: Game Playing

T = playing Checkers

P = win rate against opponents

E = playing games against itself

Example 1 : Tic-Tac-Toe

- ▶ How to program the computer to play tic-tac-toe?

		X		O		X		O		X		O		X		O		X		O		X

- ▶ Option A: The programmer writes explicit rules, e.g. 'if the opponent has two in a row, and the third is free, stop it by placing your mark there', etc (lots of work, difficult, not at all scalable!)
- ▶ Option B: Go through the game tree, choose optimally (for non-trivial games, must be combined with some heuristics to restrict tree size)
- ▶ Option C: Let the computer try out various strategies by playing against itself and others, and noting which strategies lead to winning and which to losing (='machine learning')

Example 2 : face recognition

- ▶ Face recognition is hot (facebook, apple; security; ...)
- ▶ Programmer writes rules: ~~"If short dark hair, big nose, then it is Mikko"~~ (impossible! how do we judge the size of the nose?!)
- ▶ The computer is shown many (image, name) example pairs, and the computer learns which features of the images are predictive (difficult, but not impossible)



patrik



antti



doris



patrik

...


...



?

Example 3: Prediction of search queries

- ▶ The programmer provides a standard dictionary (words and expressions change!)
- ▶ Previous search queries are used as examples!





what do **dreams** mean


what do **contractions** feel like

what do **you want** from me

what do **turtles** eat

[An Online Guide To Dream Interpretation](#)
www.dreammoods.com/ 
20 Aug 2011 – We realize that your **dreams** are unique; no other individual **can** have your ... It provides you with insight into your own self and a **means** for ...
[Dream Dictionary](#) - [Teeth Dreams](#) - [Common Dreams](#) - [Chase Dreams](#)

[Dream Moods A-Z Dream Dictionary](#)
www.dreammoods.com/dreamdictionary/ 
1 Jul 2011 – In analyzing your **dreams**, you **can** learn about your deep ...
[Type Of Dreams](#) - [Dream Themes](#) - [Your Dream Symbol Interpretation](#)

 Show more results from dreammoods.com

Example 4: Ranking search results

- ▶ Various criteria for ranking results
- ▶ What do users click on after a given search? Search engines can learn what users are looking for by collecting queries and the resulting clicks.

nokia

Noin 186 000 000 tulosta (0,08 sekuntia)

[Mukautettu](#) »

Nokia Online Kauppa

[Nokia.fi/kauppa](#) Helppoa ja sujuvaa - osta puhelin ja lisälaitteet Nokian kaupasta. Ilmainen autonavigointi ja teline - Ilmaiset karttapalvelut - Lisälaitteet - Puhelimet

Nokia, Finland - Wikipedia, the free encyclopedia ☆ - [Käännä tämä sivu]

Nokia is a town and a municipality on the banks of the Nokianvirta River (Kokemäenjoki) in the region of Pirkanmaa, some 15 kilometres (9 mi) west of ...

[en.wikipedia.org/wiki/Nokia,_Finland](#) - Välimuistissa - Samankaltaisia

Nokia - Wikipedia, the free encyclopedia ☆ - [Käännä tämä sivu]

Nokia Corporation OMX: NOK1V, NYSE: NOK, FWB: NOA3) is a Finnish ...

[en.wikipedia.org/wiki/Nokia](#) - Välimuistissa - Samankaltaisia

Nokia 5700 XpressMusic – Wikipedia ☆

Nokia 5700 XpressMusic on vuonna 2007 julkaistu nuorten musiikkipuhelin ...

[fi.wikipedia.org/wiki/Nokia_5700_XpressMusic](#) - Välimuistissa - Samankaltaisia

[+](#) Näytä lisää tuloksia kohteesta wikipedia.org

Nokia (nokia) on Twitter ☆ - [Käännä tämä sivu]

News and updates from Nokia. The main tweeps at the channels are @jussipekka & @JGallo02.

[twitter.com/nokia](#) - Välimuistissa - Samankaltaisia

Ovi Musiikki - porttisi musiikin maailmaan ☆

Aloitussivu · Nokia Ovi Player · Ovi Musiikki Unlimited Nokia.com; Copyright ©2010

Nokia. Kaikki oikeudet pidätetään.

[music.ovi.com/fi/fi/pc](#) - Välimuistissa

YouTube - Lex Nokia anti-ad 2A: "Perustuslaki" ☆

29. tammikuu 2009 ... Urkintalaki.fi:n masinoma Lex Nokia -lakiehdotuksen vastainen mainos 2a. " Perustuslaki".

[www.youtube.com/watch?v=0tDhemyzB3k](#) - Välimuistissa - Samankaltaisia

Example 5

- ▶ Self-driving cars:
 - ▶ Sensors (radars, cameras) superior to humans
 - ▶ How to make the computer react appropriately to the sensor data?

SMARTER THAN YOU THINK

Google Cars Drive Themselves, in Traffic



Ramon Romero for The New York Times

Example 6

- ▶ Machine translation:
 - ▶ Traditional approach: Dictionary and explicit grammar
 - ▶ More recently, *statistical* machine translation based on example data is increasingly being used

Google kääntäjä

Kielestä: suomi ▼ ⇅ Kielelle: englanti ▼ Käännä

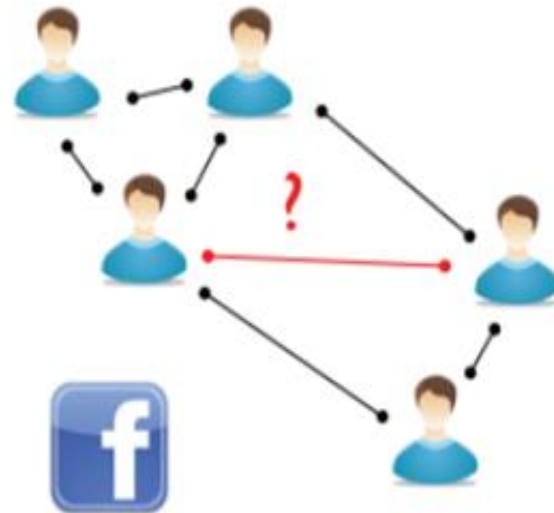
Tietojenkäsittelytieteen opinnot antavat erinomaisen pohjan työskentelylle kaikkialla, missä kehitetään tai sovelletaan tietotekniikkaa.

Käännös (suomi > englanti)

Computer studies provide an excellent foundation for the work, wherever applicable, or to develop information technology.

Example 7

- Prediction of friends in Facebook, or prediction of who you'd like to follow on Twitter.



Problem setup

Consider the case of Game playing, According to ML definition, it is a computer program that improve its **performance** on a given **task** with **experience** :

T = playing Checkers

P = win rate against opponents

E = playing games against itself



1- Task

▶ **Task:** What is the problem that the program is solving?

- Machine learning allows us to tackle tasks that are too difficult to solve with ***fixed programs written and designed by human beings.***
- Machine learning is interesting because developing our understanding of machine learning entails **developing our understanding of the principles that underlie intelligence.**

1- Task (Cont.)

Some of the most common machine learning tasks include the following:

1- Classification : In this type of task, the computer program is asked to specify which of k categories some input belongs to.

-Classification Algorithms attempt to map inputs into one of a set of classes (Colors, Good and Bad Credit Risks)

2- Regression : In this type of task, the computer program is asked to predict a numerical value given some input.

-Regression Algorithms attempt to map inputs into continuous output (Integers, Real Numbers, Vectors, etc.)

1- Task (Cont.)

Some of the most common machine learning tasks include the following:

3- Transcription: In this type of task, the machine learning system is asked to observe a relatively unstructured representation of some kind of data and transcribe it into discrete, textual form.

E.g.: optical character recognition (OCR), where the computer program is shown a photograph containing an image of text and is asked to return this text in the form of a sequence of characters.

4- Machine translation: In a machine translation task, the input already consists of a sequence of symbols in some language, and the computer program must convert this into a sequence of symbols in another language.

2- Experience

► **Experience:** What is the data (examples) that the program is using to improve its performance?

- **Experience** is a dataset.
- **A dataset** is a collection of many examples.
- **An example** is a collection of **features** that have been quantitatively measured from some object or event that we want the machine learning system to process.

3- performance level

- ▶ **Performance measure:** How is the performance of the program (when solving the given task) evaluated?

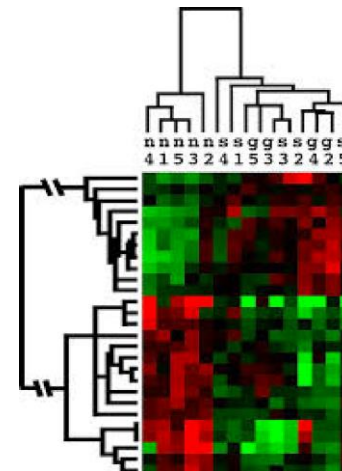
In order to evaluate the abilities of a machine learning algorithm, we must design a quantitative measure of its performance. Usually this performance measure P is specific to the task T being carried out by the system.

- Accuracy : Accuracy is just the proportion of examples for which the model *produces the correct output*.
- Error Rate : the proportion of examples for which the model *produces an incorrect output*.

WHEN DO WE USE MACHINE LEARNING?

ML is used when:

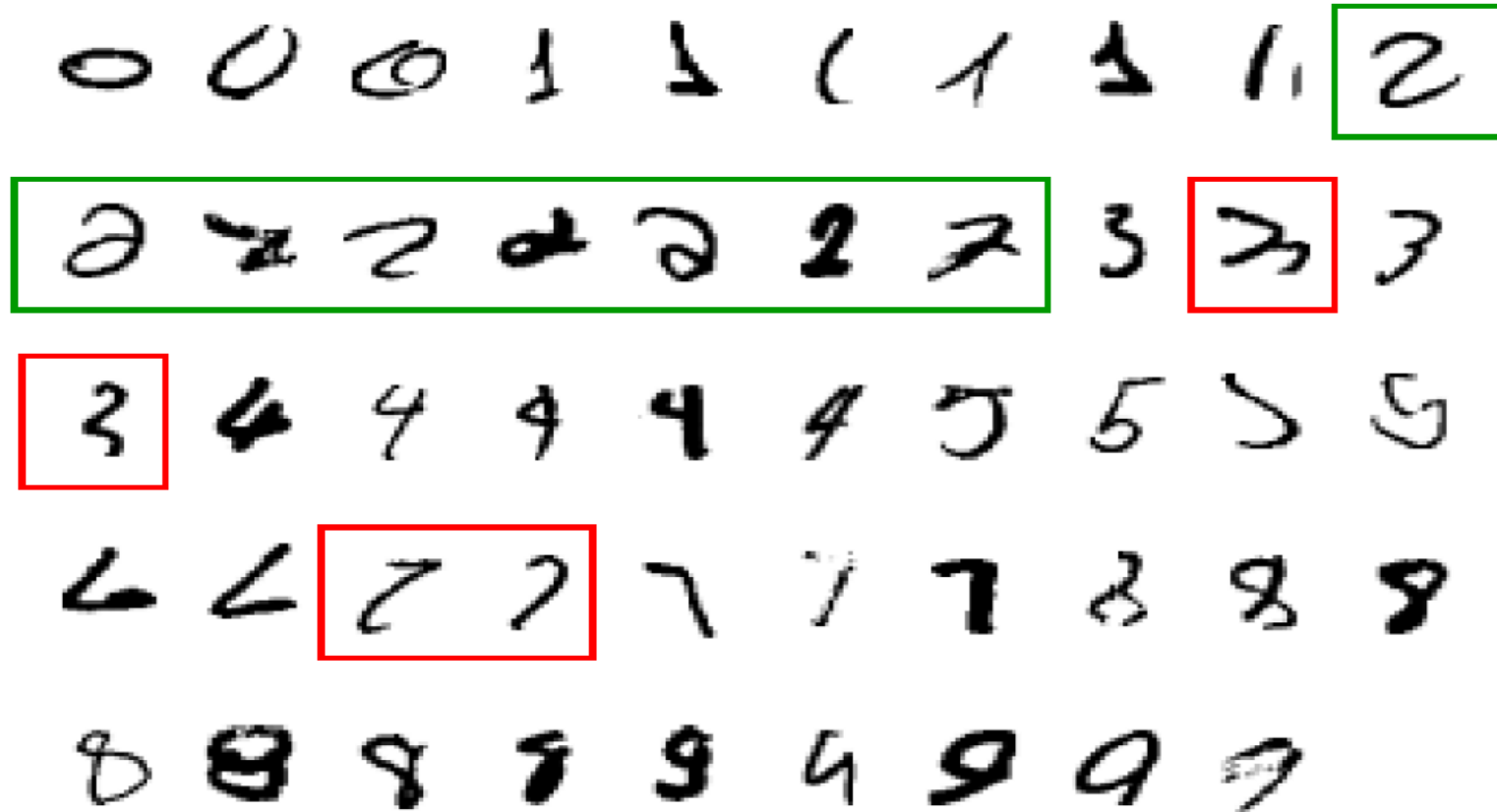
- ☐ • Human expertise does not exist (navigating on Mars)
- ☐ • Humans can't explain their expertise (speech recognition)
- ☐ • Models must be customized (personalized medicine)
- ☐ • Models are based on huge amounts of data (genomics)



WHEN DO WE USE MACHINE LEARNING?

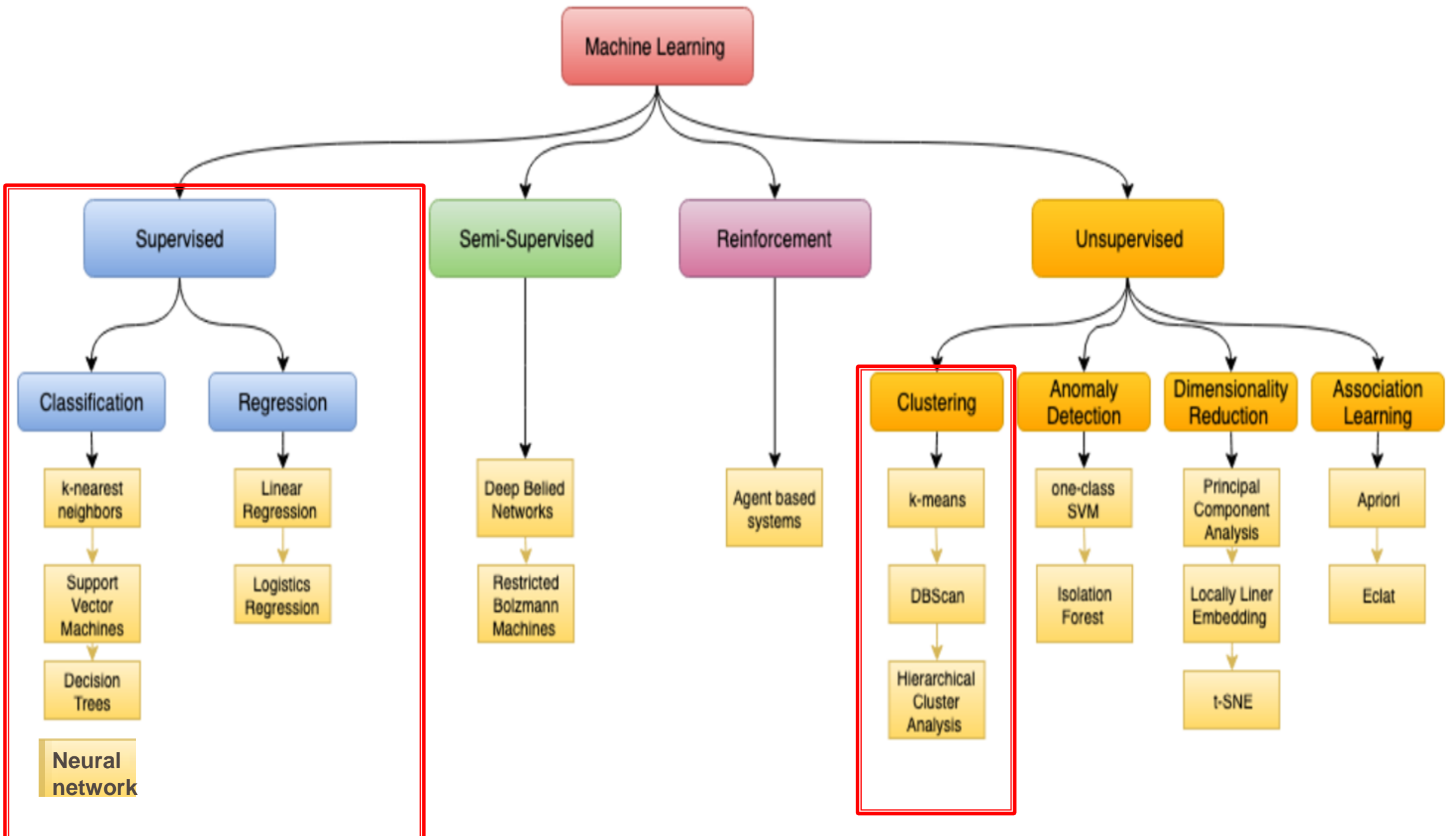
A classic example of a task that requires machine learning:

It is very hard to say what makes a 2



Some more examples of tasks that are best solved by using a learning algorithm

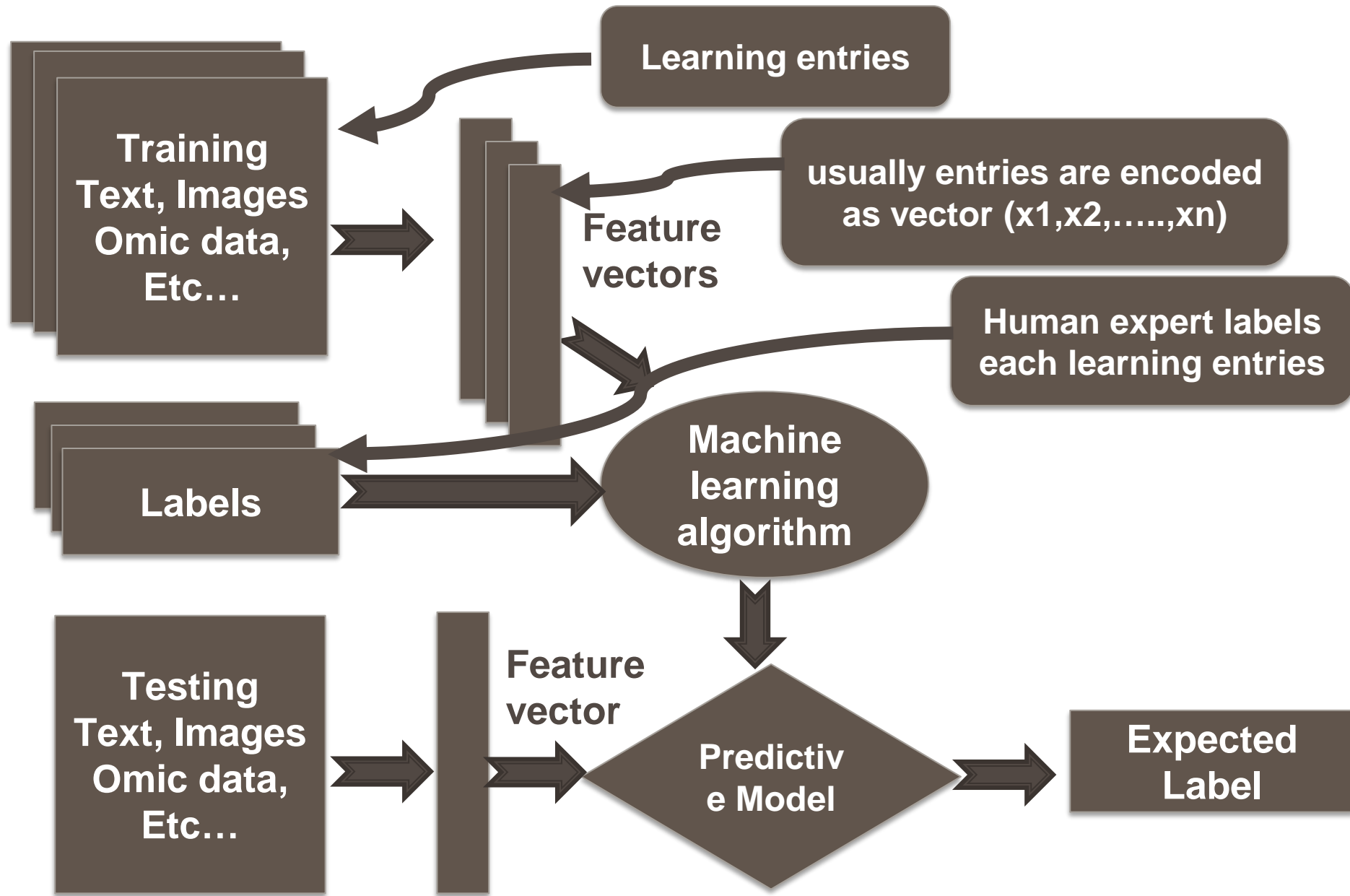
- Recognizing patterns:
 - Facial identities or facial expressions
 - Handwritten or spoken words
 - Medical images
- Generating patterns:
 - Generating images or motion sequences
- Recognizing anomalies:
 - Unusual credit card transactions
 - Unusual patterns of sensor readings in a nuclear power plant
- Prediction:
 - Future stock prices or currency exchange rates



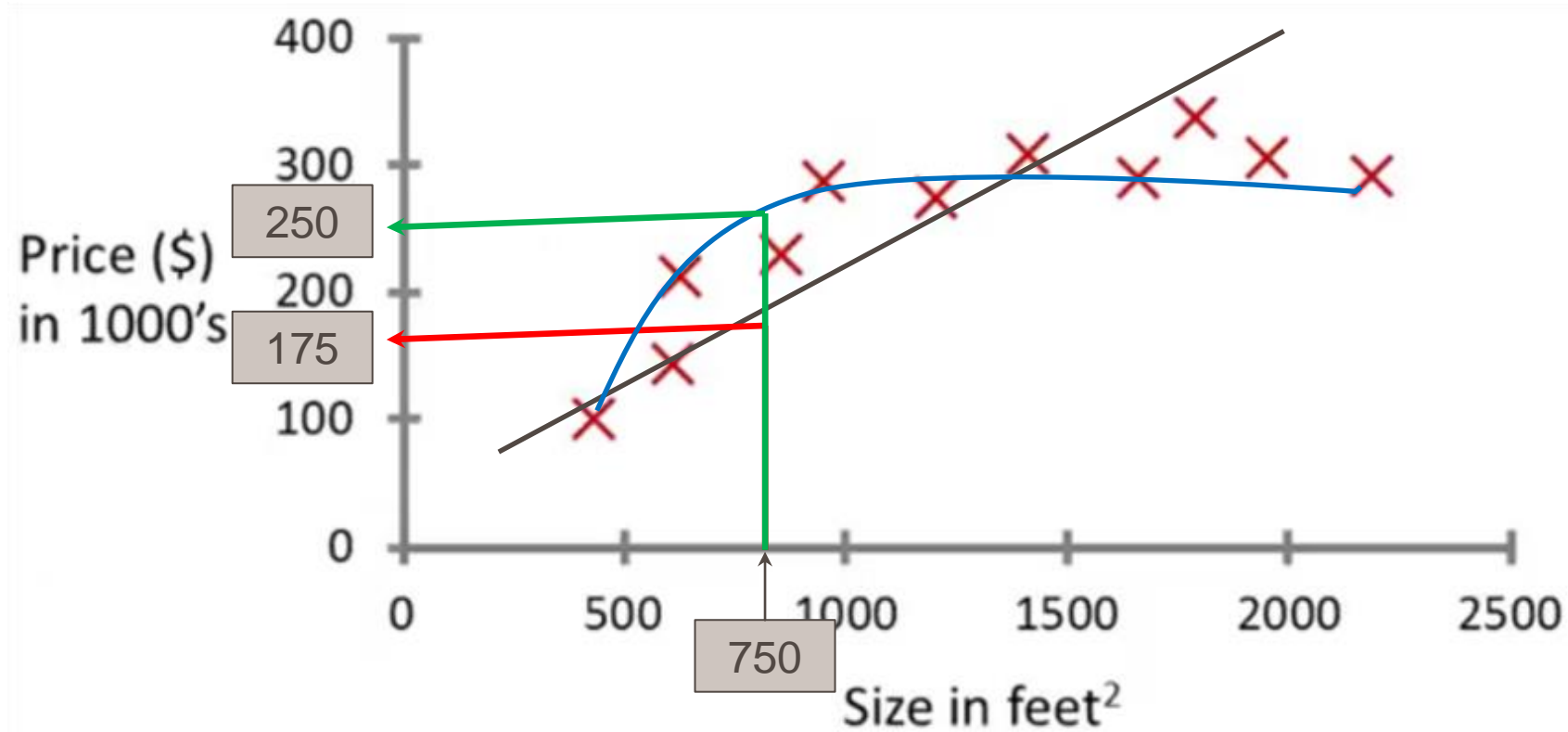
Supervised Learning

- ❑ In supervised learning, we are given a data set and already know what our correct output should look like, having the idea that there is a relationship between the input and the output.
- ❑ Supervised learning problems are categorized into "regression" and "classification" problems.
- ❑ In a regression problem, we are trying to predict results within a continuous output, meaning that we are trying to map input variables to some continuous function.
- ❑ In a classification problem, we are instead trying to predict results in a discrete output. In other words, we are trying to map input variables into discrete categories.

SUPERVISED LEARNING PARADIGM



Housing price prediction



Supervised Learning

“right answers” or “Labeled data”

given

Regression:

Predict continuous valued output

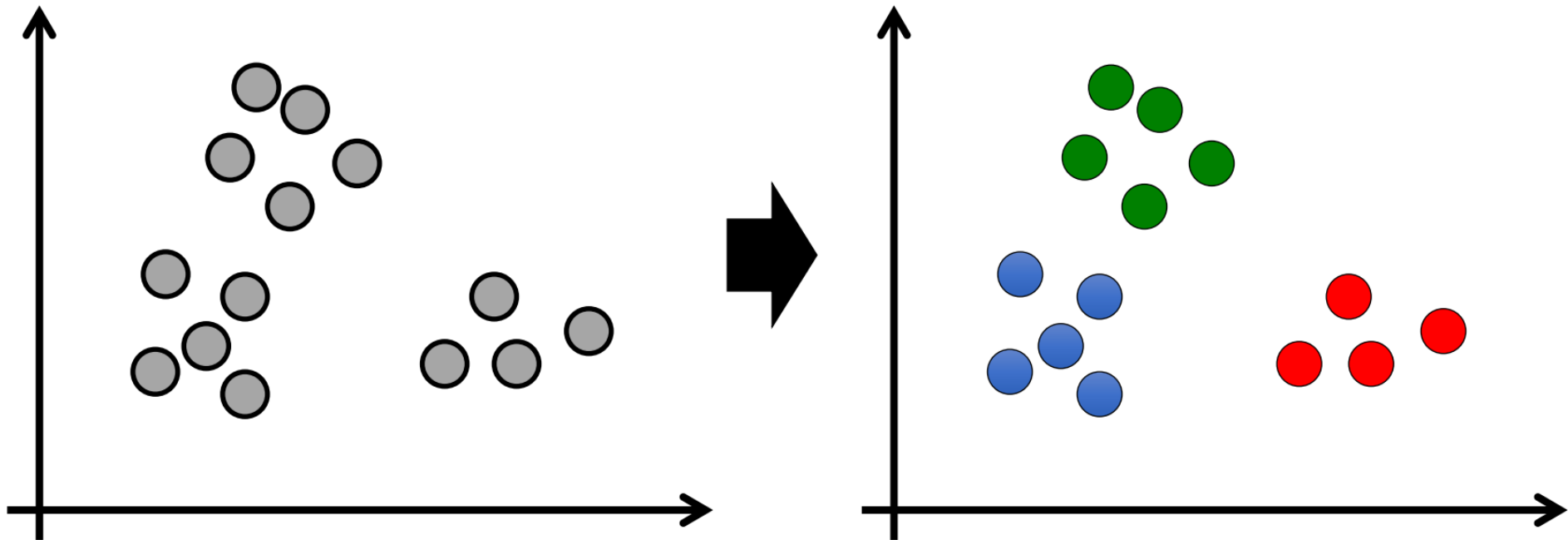
(price)

UNSUPERVISED LEARNING

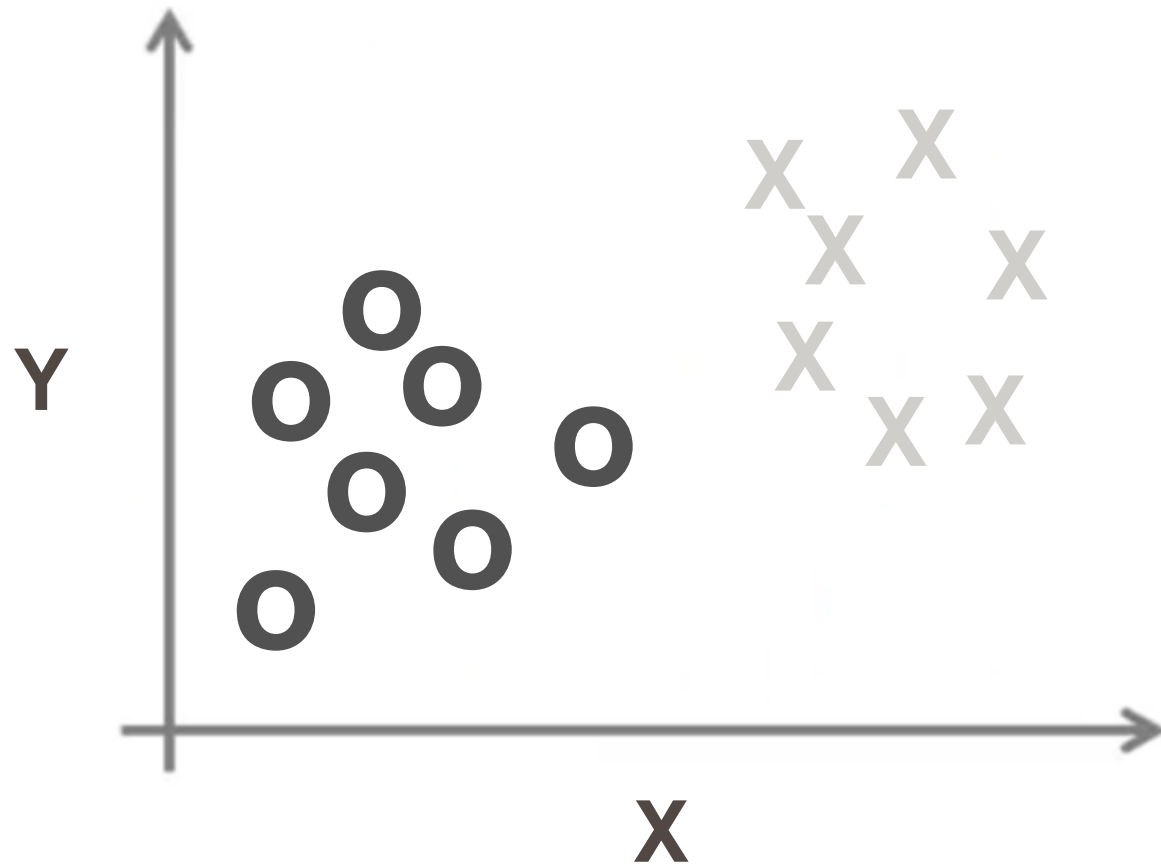
- ❑ Unsupervised learning, on the other hand, allows us to approach problems with little or no idea what our results should look like.
- ❑ We can derive structure from data where we don't necessarily know the effect of the variables.
- ❑ We can derive this structure by clustering the data based on relationships among the variables in the data.
- ❑ With unsupervised learning there is no feedback based on the prediction results, i.e., there is no teacher to correct you.

Unsupervised Learning

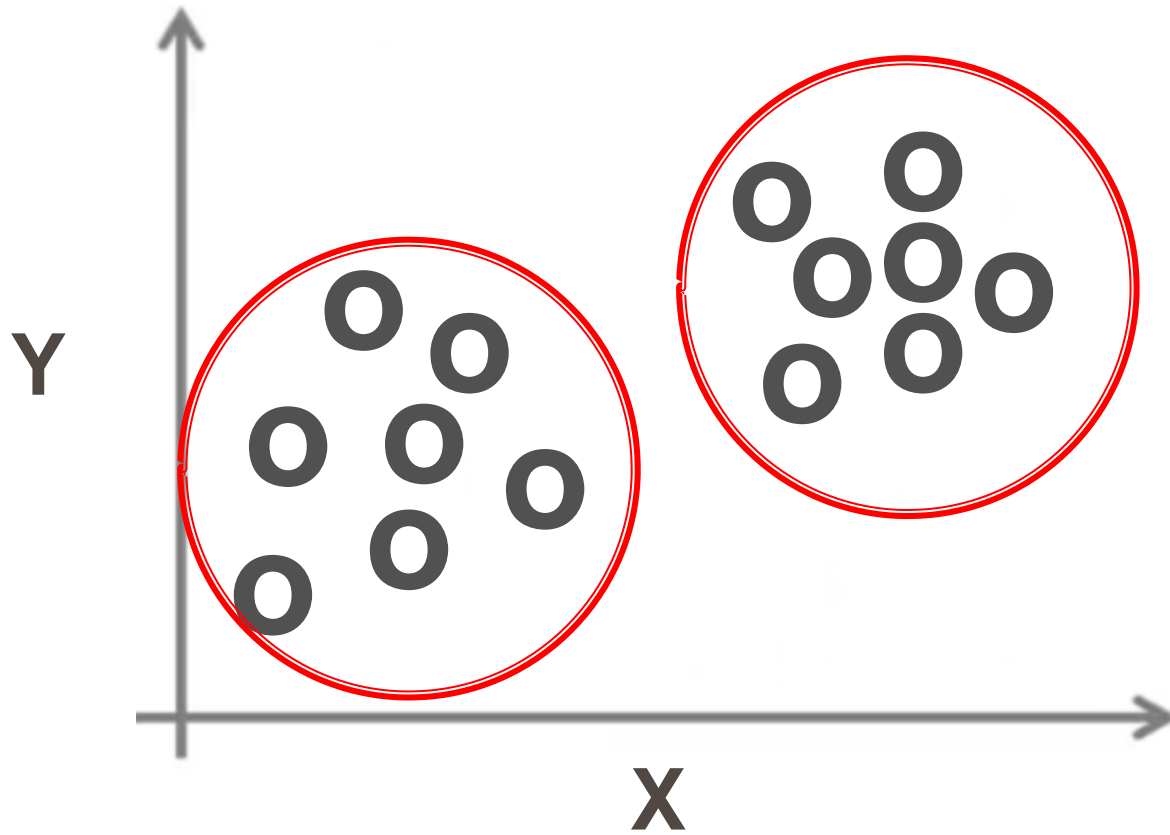
- Given x_1, \dots, x_n (no labels), output hidden structure in x 's
 - E.g., clustering



SUPERVISED

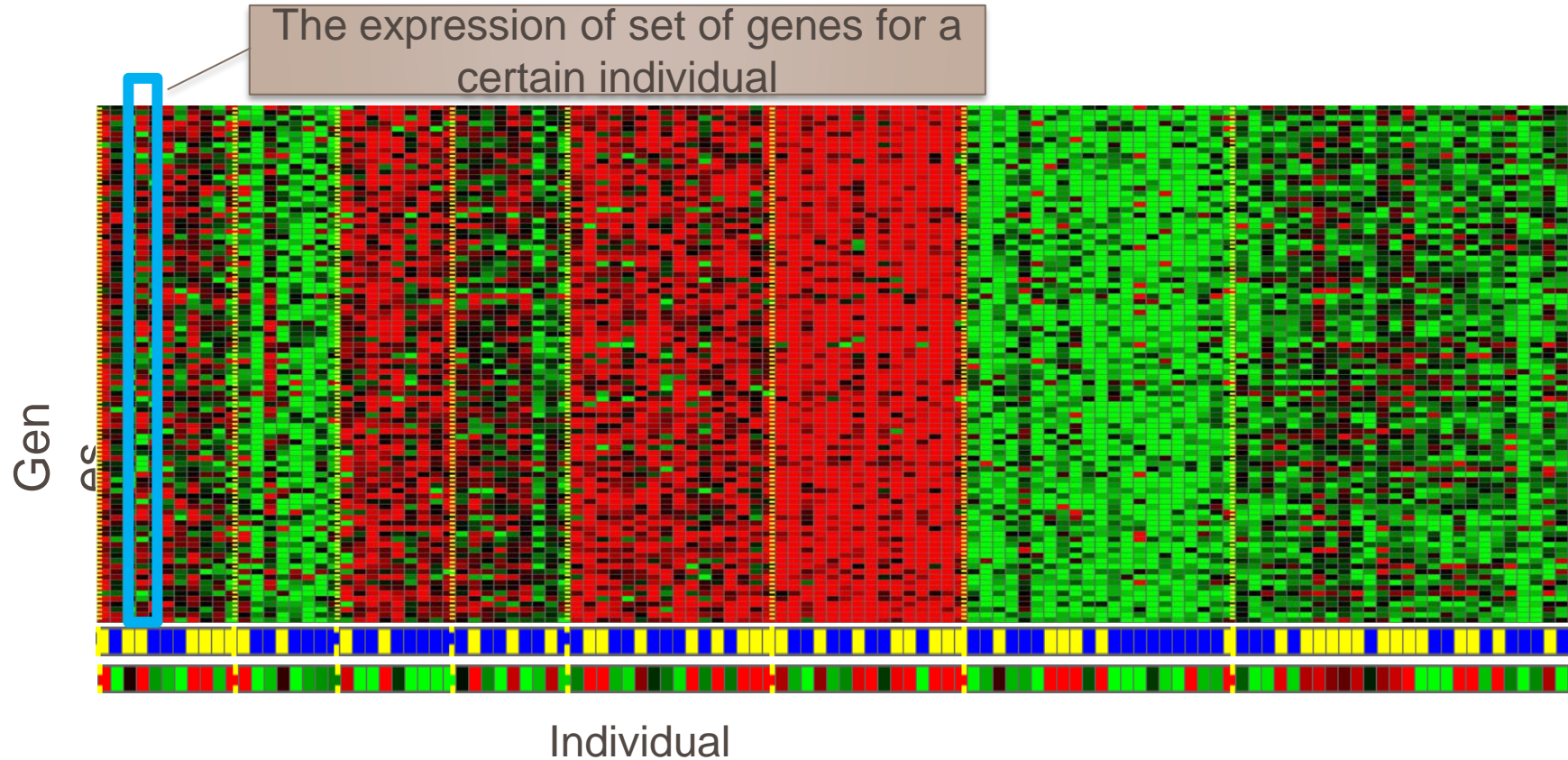


Unsupervised



Unsupervised learning is where you only have input data and **no corresponding output variables**.

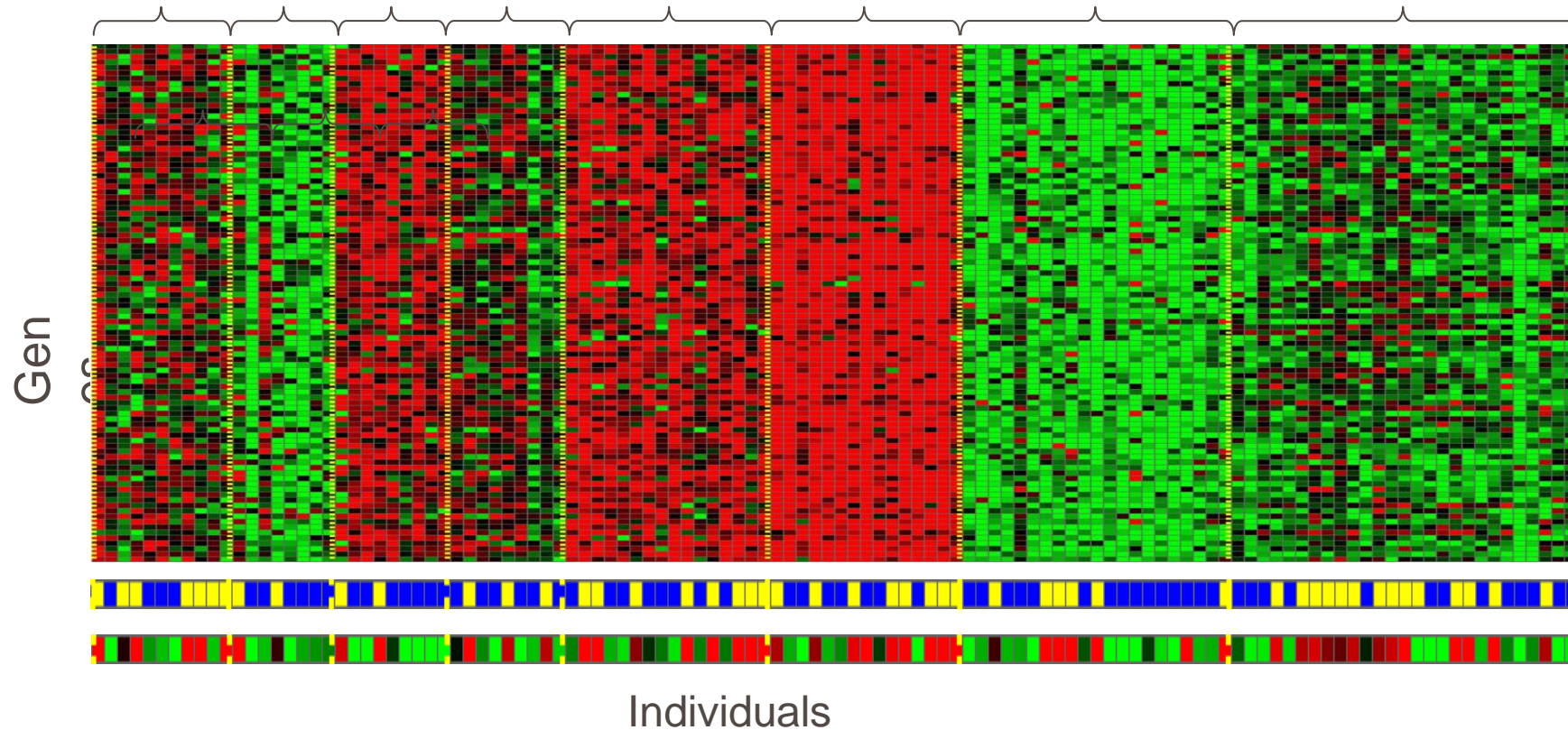
Unsupervised Learning applications



We need to cluster the individuals based of the similarity of their genes.



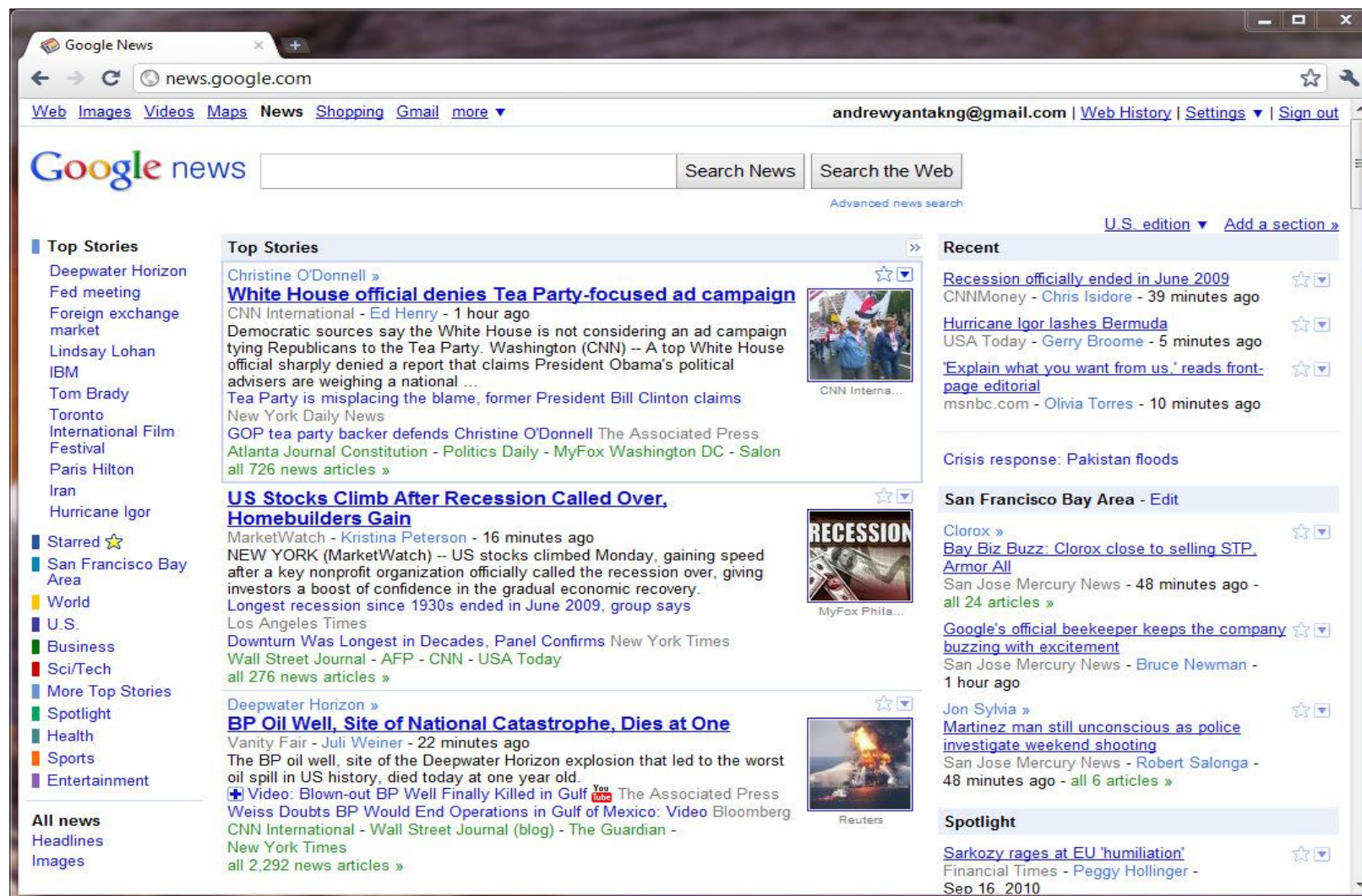
Unsupervised Learning applications



Microarray data

- Have a group of individuals
- On each measure expression of a gene
- Run algorithm to cluster individuals into types of people

Unsupervised Learning applications



Google News

news.google.com

Web Images Videos Maps News Shopping Gmail more

Google news

Search News Search the Web

Adjusted news feed

U.S. edition Add a section

Top Stories

- Deepwater Horizon
- Fed meeting
- Foreign exchange market
- Lindsay Lohan
- IBM
- Tech Grady
- Turkey
- International Film Festival
- Plans for Iran
- Hurricane Igor

Top Stories

Chrisine O'Donnell

Whyte House official denies Tea Party-focused ad campaign

Chrisine O'Donnell • 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 came President Obama's political advisers are weighing a national

Tea Party is misrepresing the Marine, former President Bill Clinton claims

New York Daily News

GOP tea party leader defends Chrisine O'Donnell The Associated Press

Atlanta Journal-Constitution • Politico Daily • MyFox Washington DC • Salon

all 726 news articles

US Stocks Climb After Recession Called Over, Homebuilders Gain

MarketWatch • Kristina Peterson • 18 minutes ago

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

Largest recession since 1930s ended in June 2009, group says

Los Angeles Times

Downton Was Longest in Decades, Panel Confirms New York Times

Wall Street Journal • AP • CNN • USA Today

all 10 news articles

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

Vanity Fair • Jill Wagner • 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: Deepwater BP Well Finally Killed in Gulf

The Associated Press

Wires Doubts BP Would End Operations in Gulf of Mexico Video: Deepwater BP Well Finally Killed in Gulf

CNN International • Wall Street Journal (Blog) • The Guardian

New York Times

all 212 news articles

Recent

Recession officially ended in June 2009

ChildArmy • Chris Isidore • 30 minutes ago

Hurricane Igor lashes Bermuda

USA Today • Gerry Sheskin • 4 minutes ago

Explain what you want from us, reads bank

1925 effort

metacritic.com • Olivia Torres • 10 minutes ago

Crisis response: Pakistan floods

San Francisco Bay Area • E&E

Chase •

Ray, the Mayor, closes down to asking STOP

Donor Aid

San Jose Mercury News • all minutes ago

all 24 articles

Deepwater Horizon: BP's official health report shows the company's

buting commitment

San Jose Mercury News • Chris Newman •

hour ago

Jon Sytko •

Marine man still unconscious as police

investigate weekend shooting

San Jose Mercury News • Robert Saltsky •

48 minutes ago • all 6 articles

Spotlight

Scholarly origins of CIA headquarters

Financial Times • Peggy Hocking •

minutes ago

THE WALL STREET JOURNAL

THE SOURCE

Log In Register For Free Subscribe Now Get 2 Weeks Free

Financial Services Transport Leisure Healthcare Oil & Gas Sport Credit on the Web Housing Technology

September 20, 2010 12:10 PM EDT

BP Kills Macondo, But Its Legacy Lives On

Search The Source

Article Comments (2)

THE SOURCE: KYLE PETERSON

By James Harkin

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

Free book insurance covers killed the Macondo rig's BP oil rig photo of

AP/21, 2010

About The Source

The SOURCE is WSJ.com's business & finance blog. It is aimed to inform, educate, and

Most Recent

Articles Comments

1. Who Needs Plans II Anyway?
2. NYU Books Be Forced to Split Retail And Drinking Areas?
3. Thinking of Ramping Award Accepting?
4. BP Kills Macondo, But Its Legacy Lives On
5. My Macondo Wind a Survival to Buy It?

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

By the CNN Wire Staff

September 20, 2010 — Updated 11:11 GMT (2:17 AM)

Click to play

What next for Gulf oil spill?

STORY HIGHLIGHTS

(CNN) — The ruptured Macondo well, a mile under the Gulf of

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

Juha Hakola

guardian.co.uk, Monday 20 September 2010 04:30 GMT

Article history

BP's costs for the Deepwater Horizon disaster have hit \$10bn. Photograph: Getty Images

Learning Types

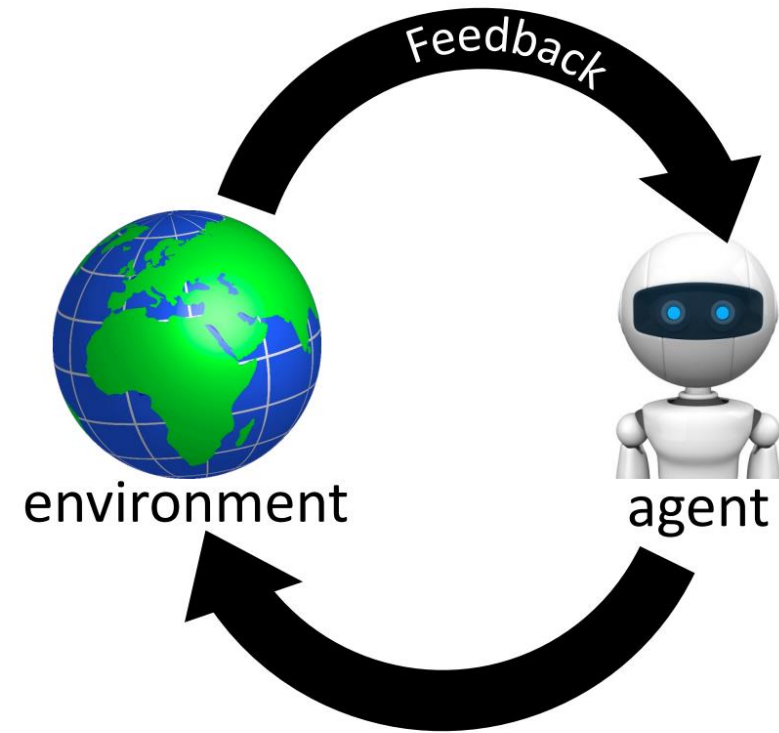
	Unsupervised	Supervised
Discrete	Clustering	Classification
Continuous	Dimensionality reduction	Regression

Machine Learning types

- Semi-supervised learning
 - mix of Supervised and Unsupervised learning
 - usually small part of data is labelled
- Reinforcement learning **Learning by interaction**
 - Model learns from a series of actions by maximizing a reward function
 - The reward function can either be maximized by penalizing bad actions and/or rewarding good actions

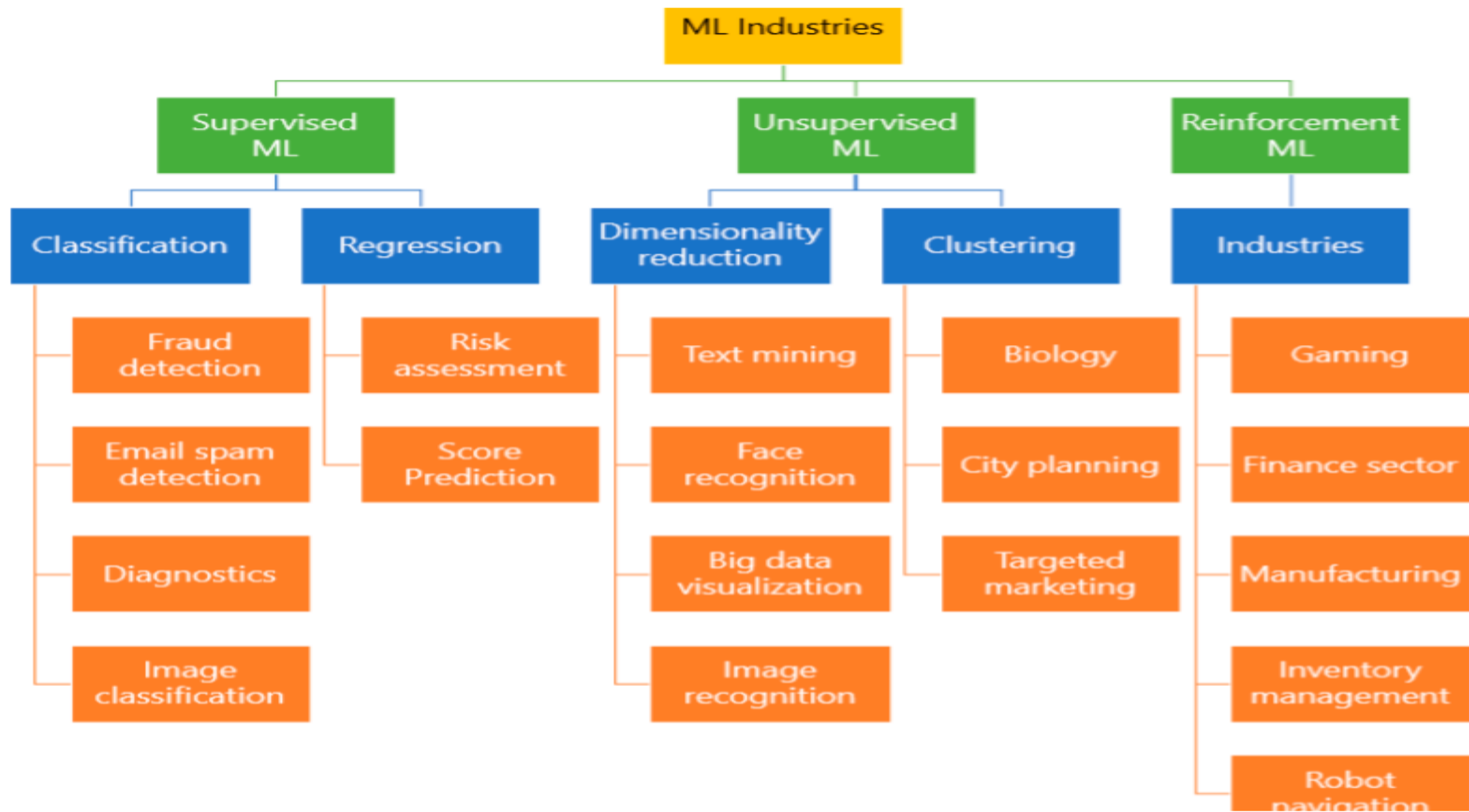
Reinforcement Learning

- Learn how to perform a task from interactions with the **environment**
- **Examples:**
 - Playing chess (interact with the game)
 - Robot grasping an object (interact with the object/real world)
 - Optimize inventory allocations (interact with the inventory system)
 - Self driven car



ML in Practice

- Understanding domain, prior knowledge, and goals
- Data integration, selection, cleaning, pre-processing, etc.
- Learning models
- Interpreting results
- Consolidating and deploying discovered knowledge
- Loop



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.

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.