

Machine learning

Presented by : Dr. Hanaa Bayomi



Lecture 1 : Introduction

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.*

What is machine learning?

➤ Definition:

Machine = computer, computer program (in this course)

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

➤ 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!

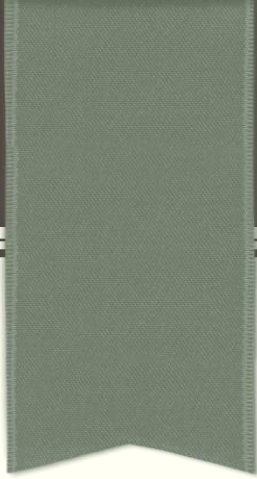
INTRODUCTION

Traditional Programming



Machine Learning





EXAMPLES

Example 1 : Tic-Tac-Toe

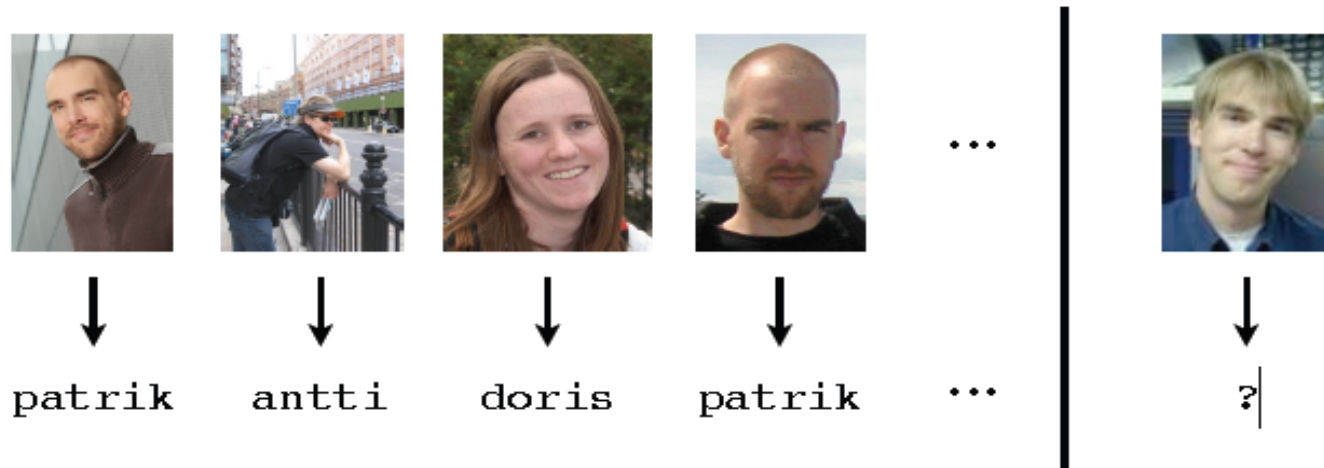
- ▶ How to program the computer to play tic-tac-toe?

		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')

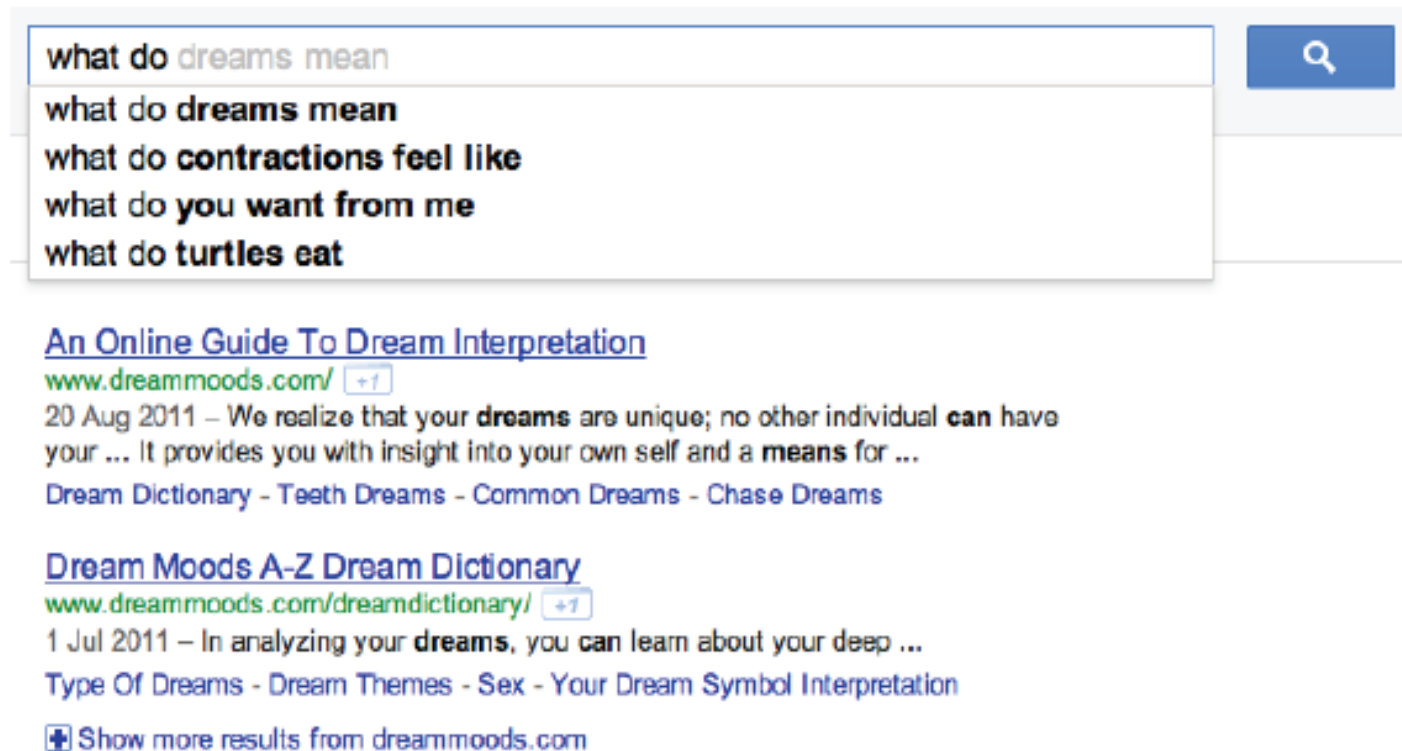
Example 2 : face recognition

- ▶ 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)



Example 3: Prediction of search queries

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



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 masinoima 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



Ramin Rohman 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

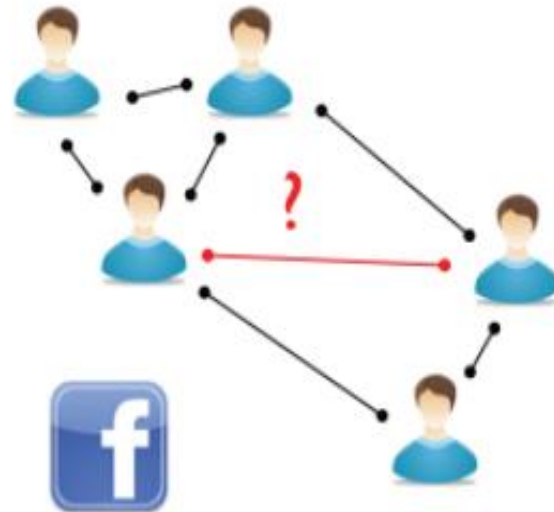
Tietojenkäsittelytieteen opinnot antavat erinomaisen pohjan työskentelylle kaikkialla, missä kehitetään tai sovelletaan tietotekniikkaa.

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

- ▶ One definition of machine learning: A computer program improves its performance on a given task with experience (i.e. examples, data).



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?

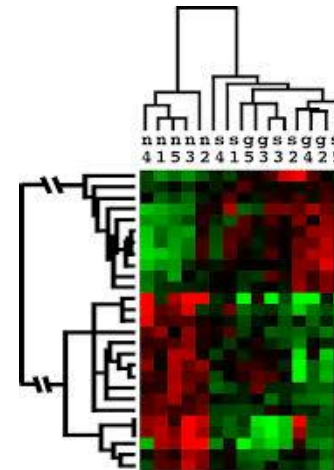
In order to evaluate the abilities of a machine learning algorithm, we must design a quantitative measure of its performance. Usually this performance measure P is specific to the task T being carried out by the system.

- 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:

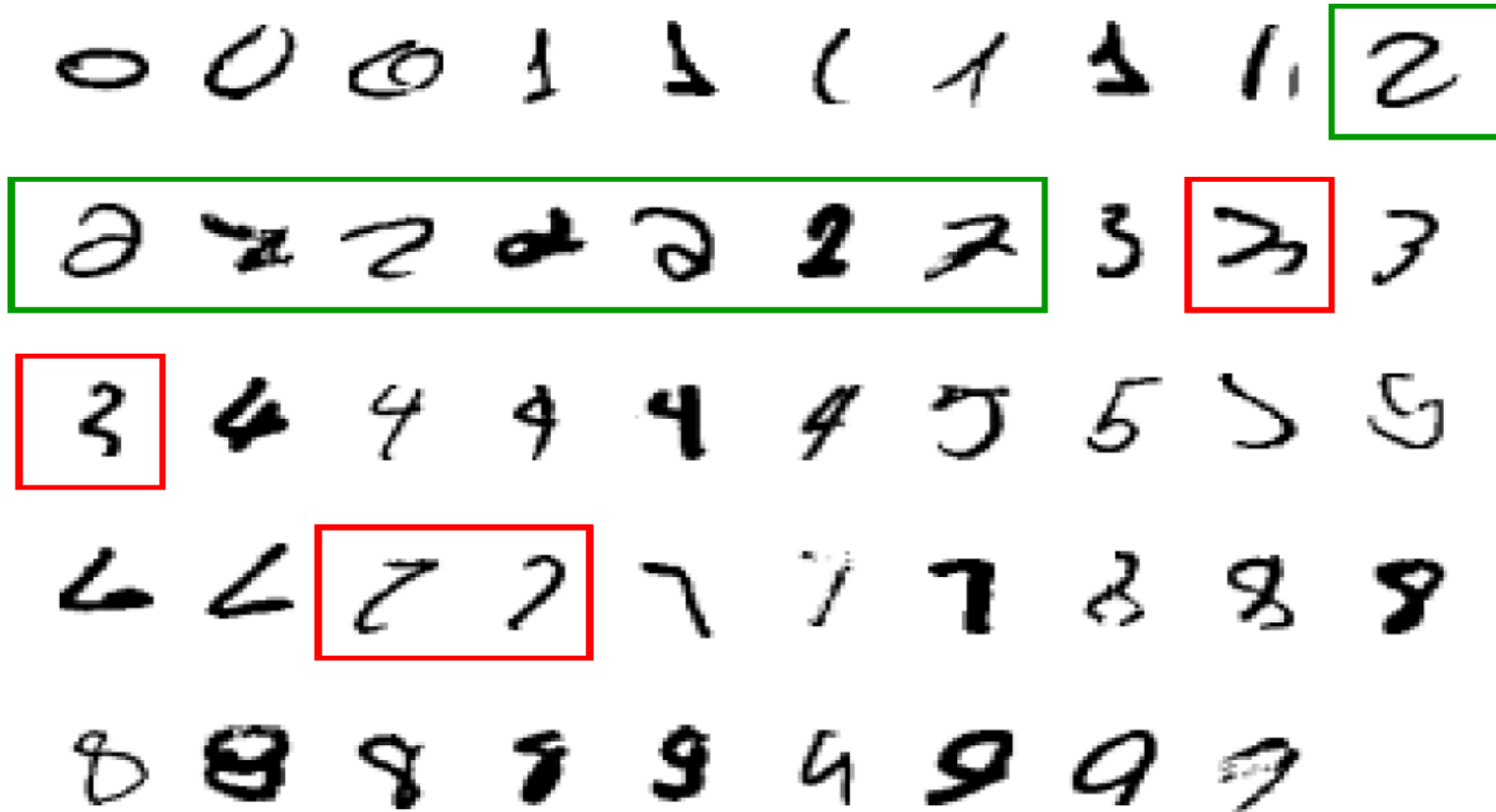
- ❑• Human expertise does not exist (navigating on Mars)
- ❑• Humans can't explain their expertise (speech recognition)
- ❑• Models must be customized (personalized medicine)
- ❑• Models are based on huge amounts of data (genomics)



WHEN DO WE USE MACHINE LEARNING?

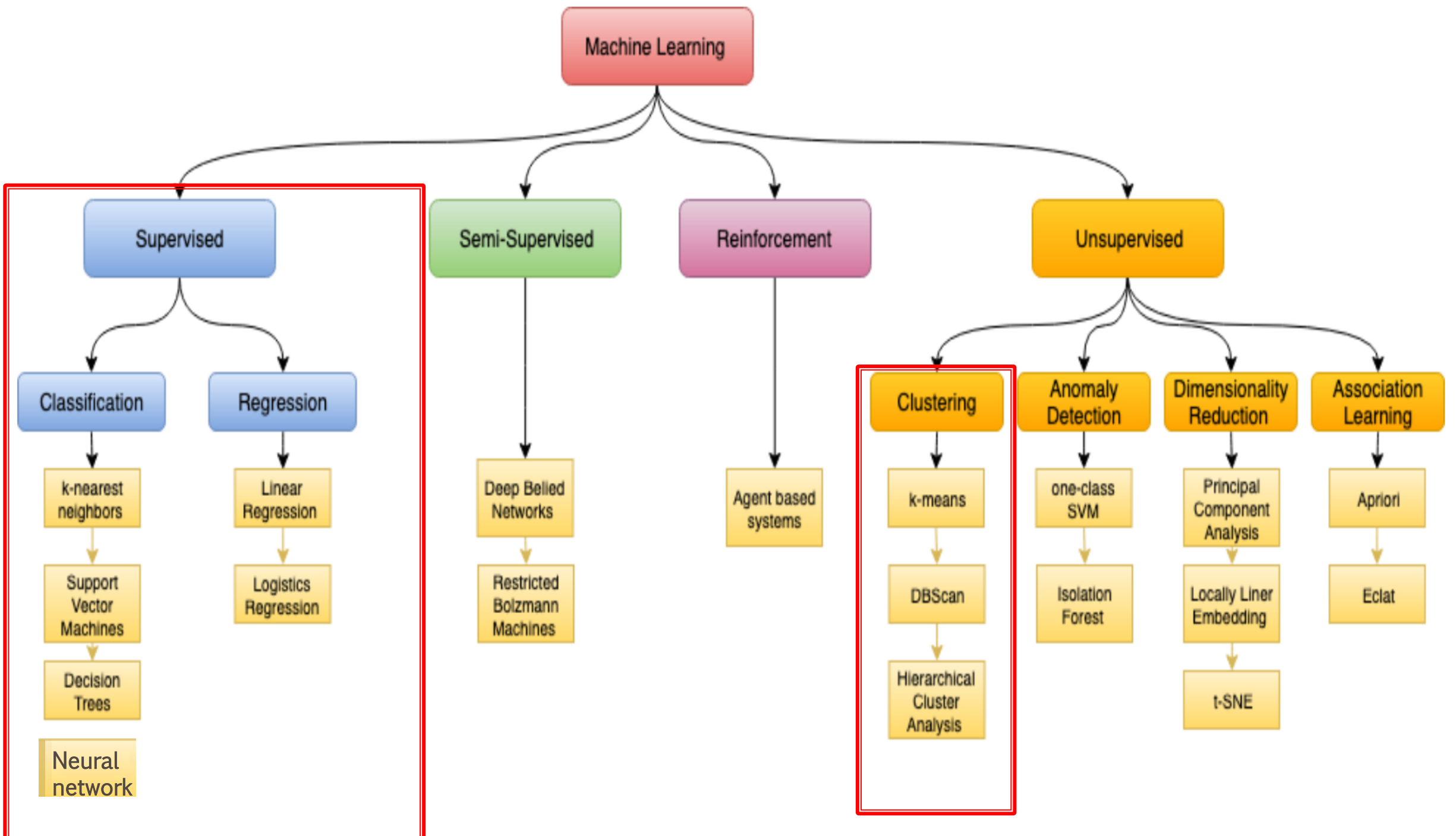
A classic example of a task that requires machine learning:

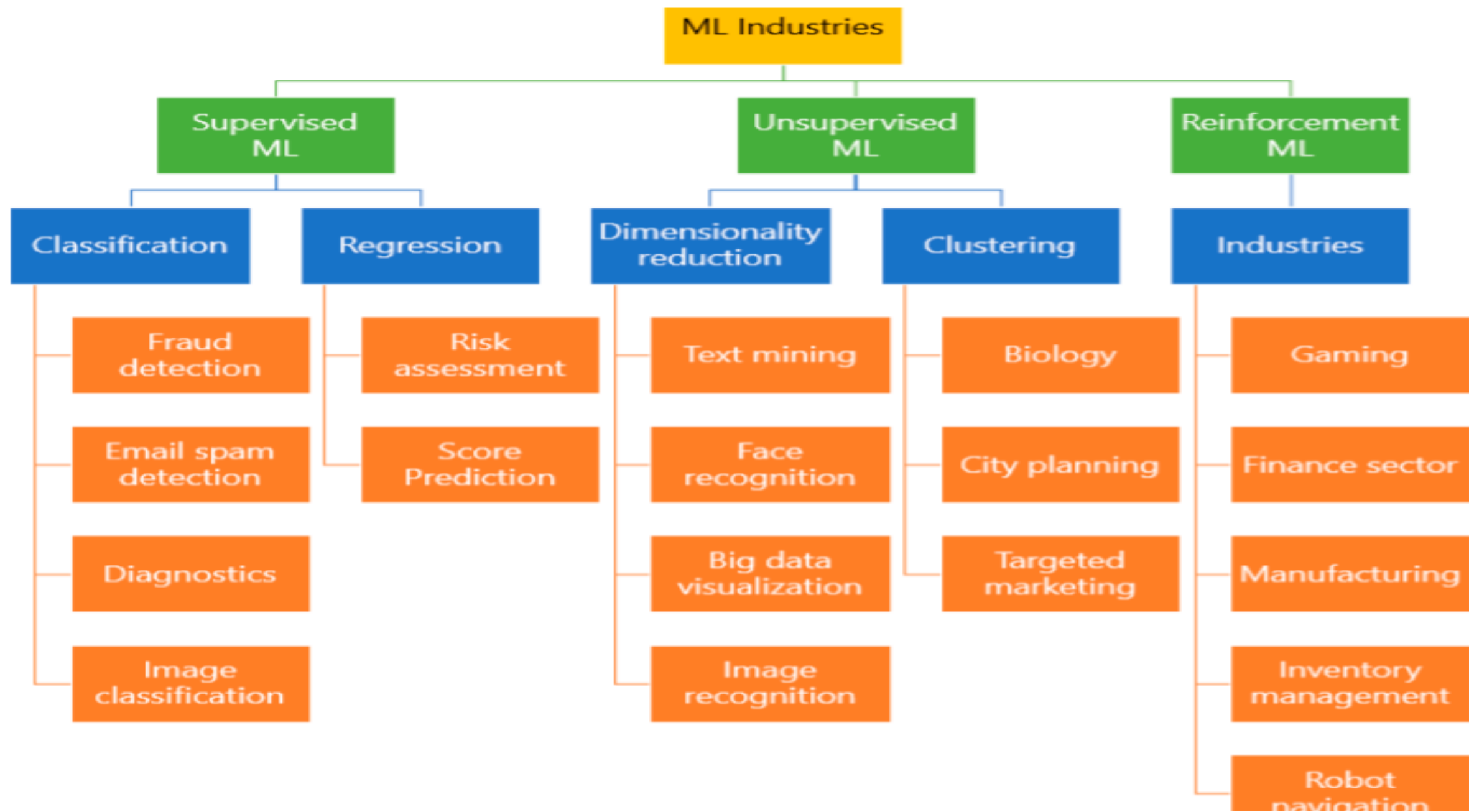
It is very hard to say what makes a 2



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

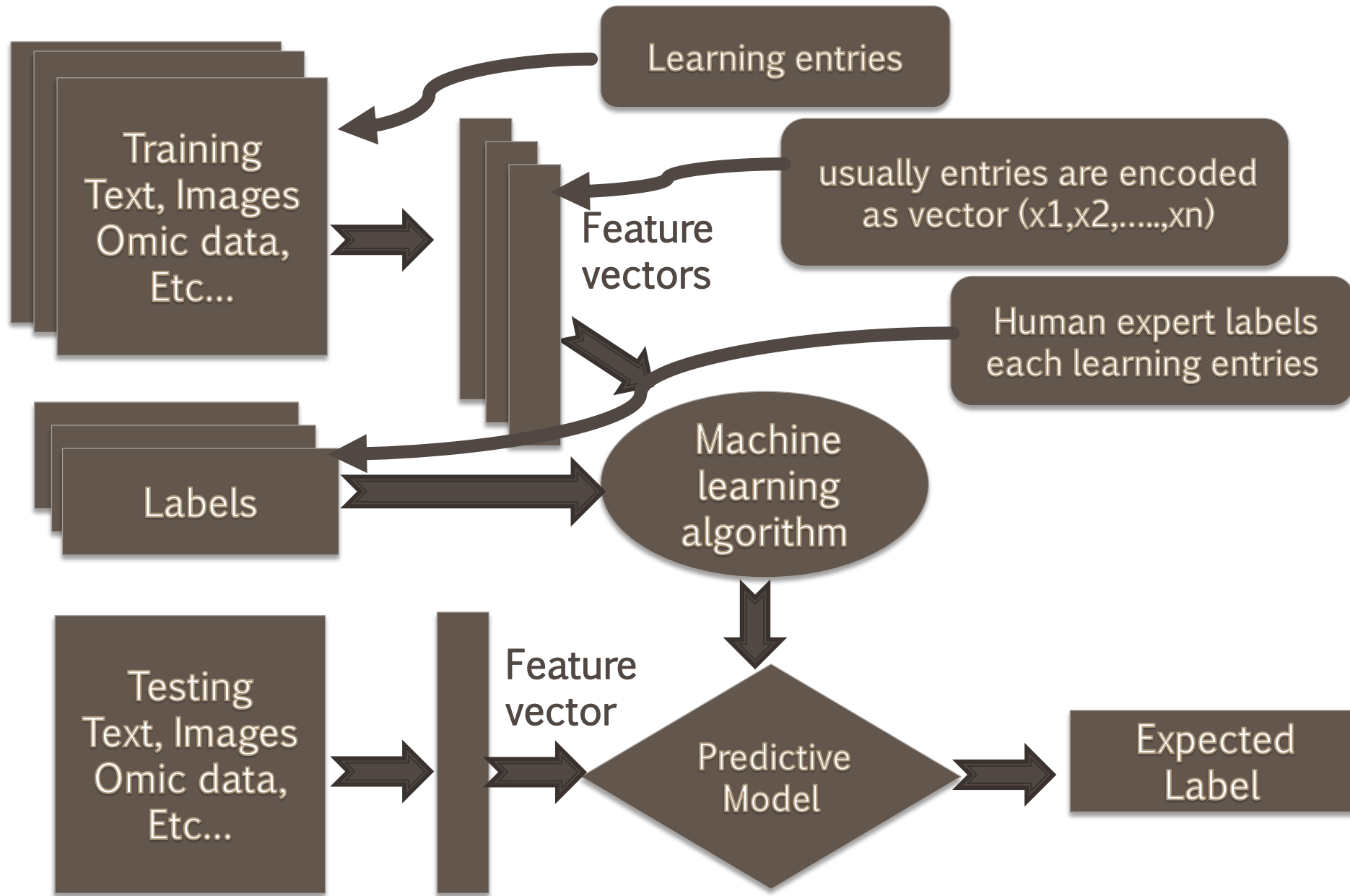




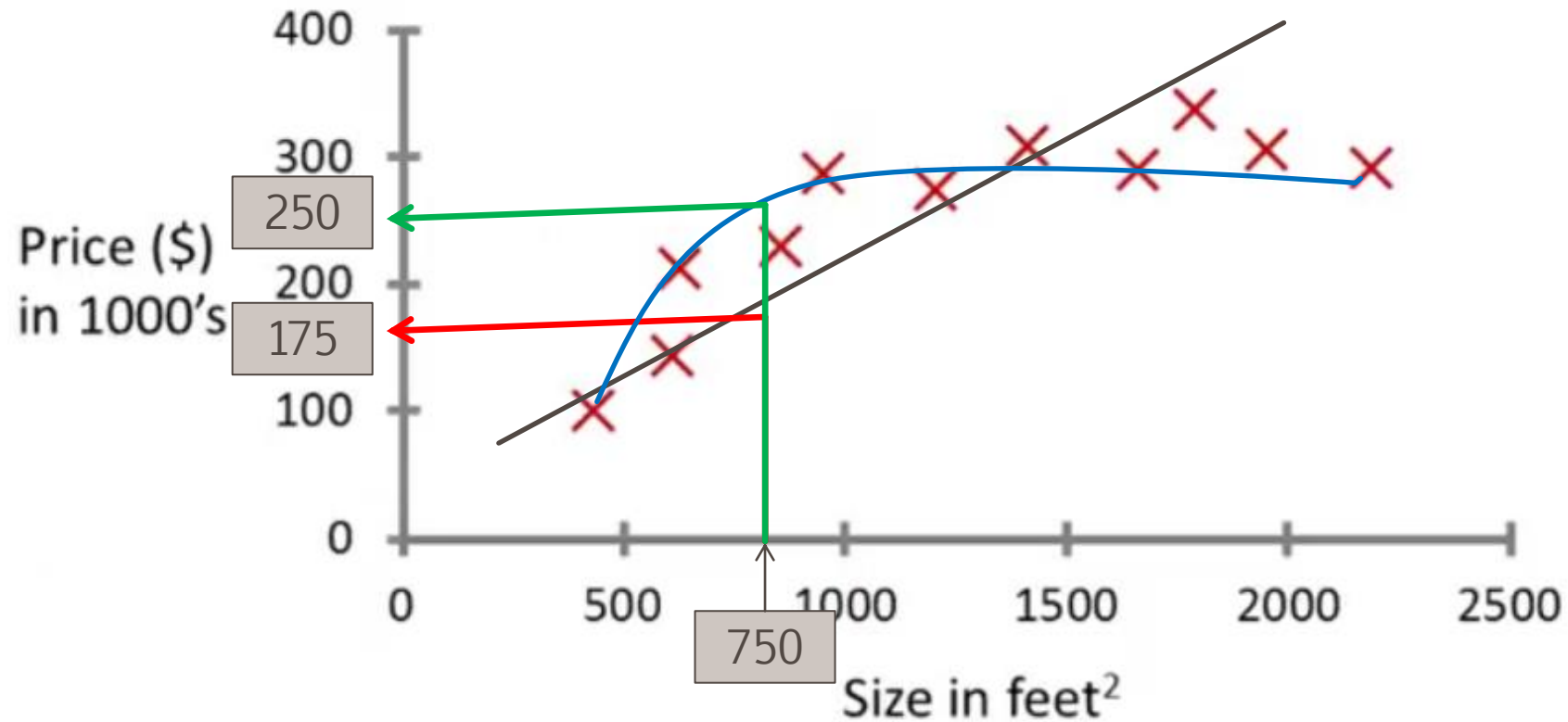
Supervised Learning

- ❑ 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

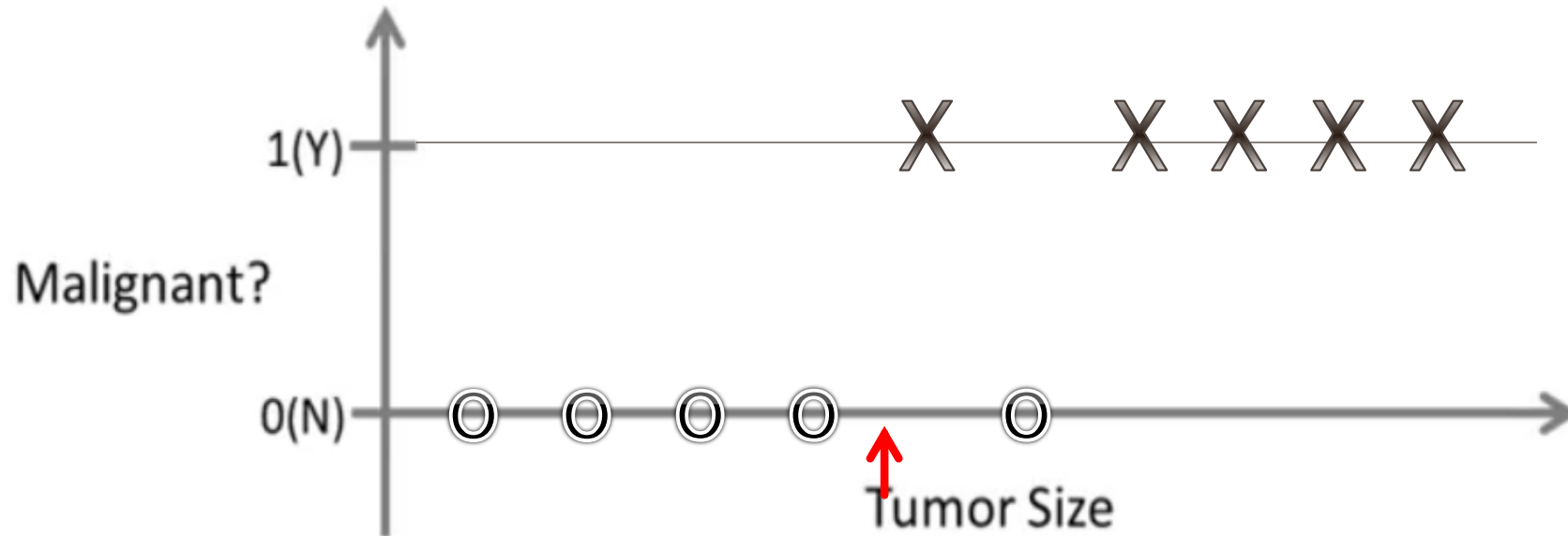
“right answers” or “Labeled data”
given

Regression:

Predict continuous valued output
(price)

Lung Cancer

(Malignant or benign)

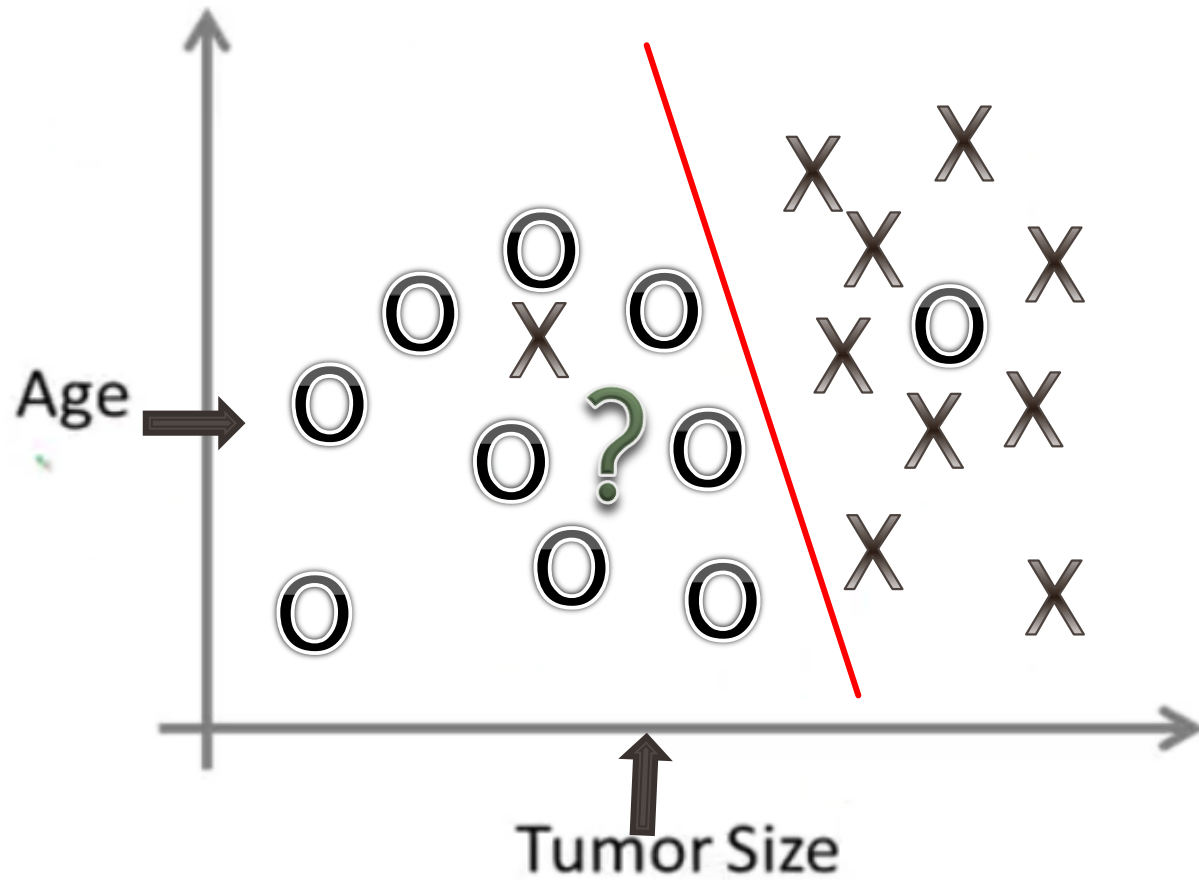


Classification

Discrete valued output (0 or 1)

Lung Cancer

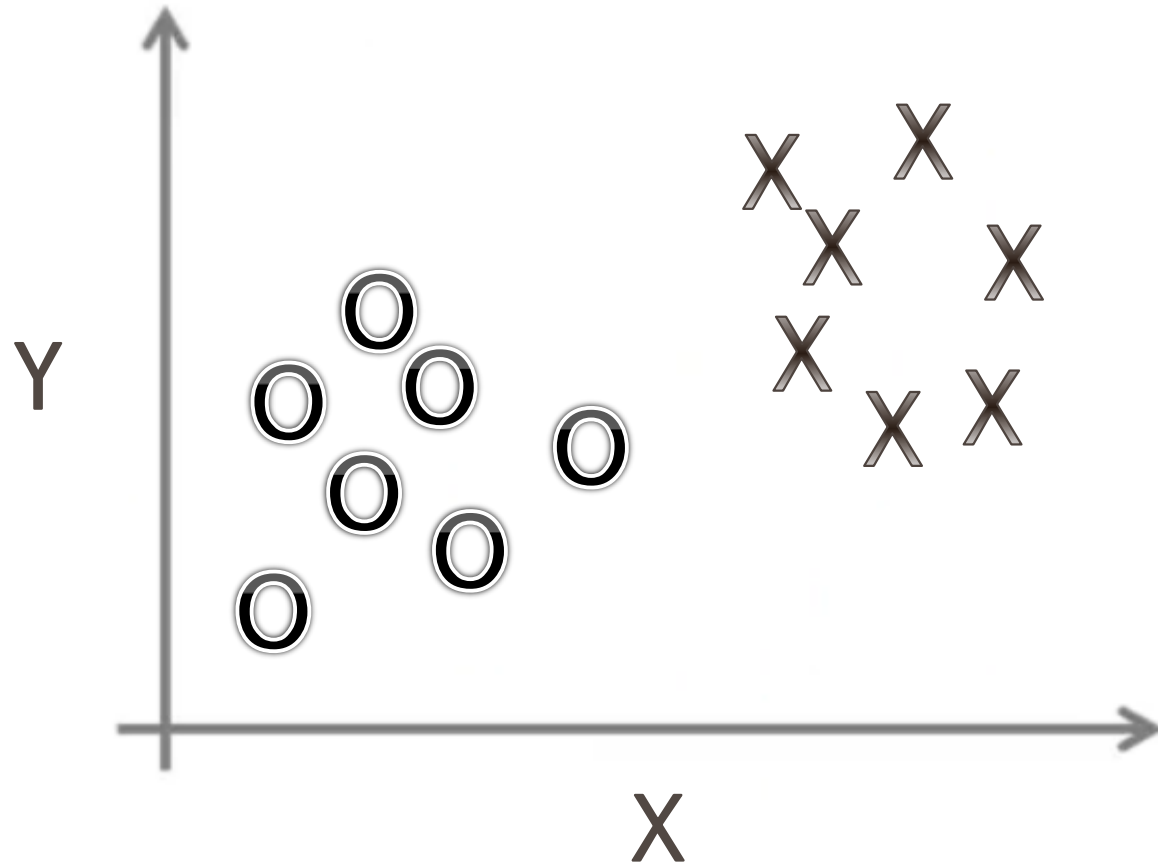
(Malignant or benign)



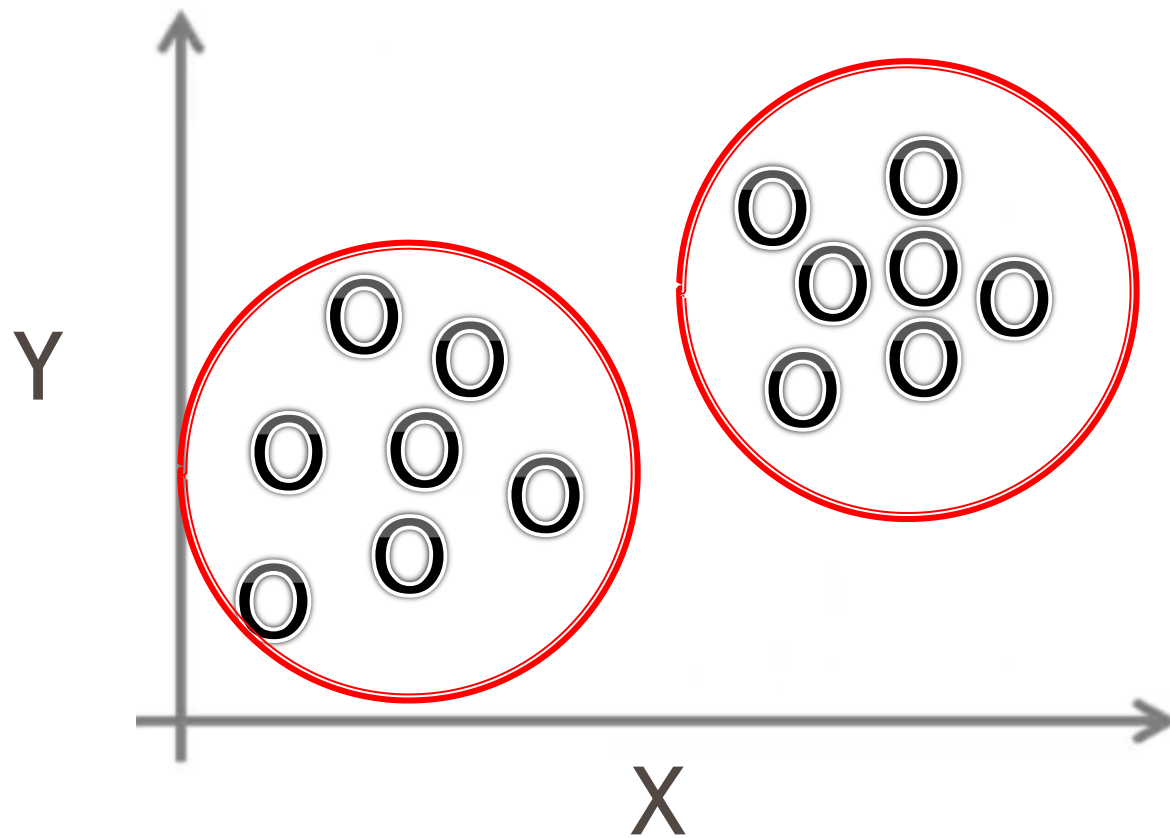
UNSUPERVISED LEARNING

- ❑ 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.

SUPERVISED

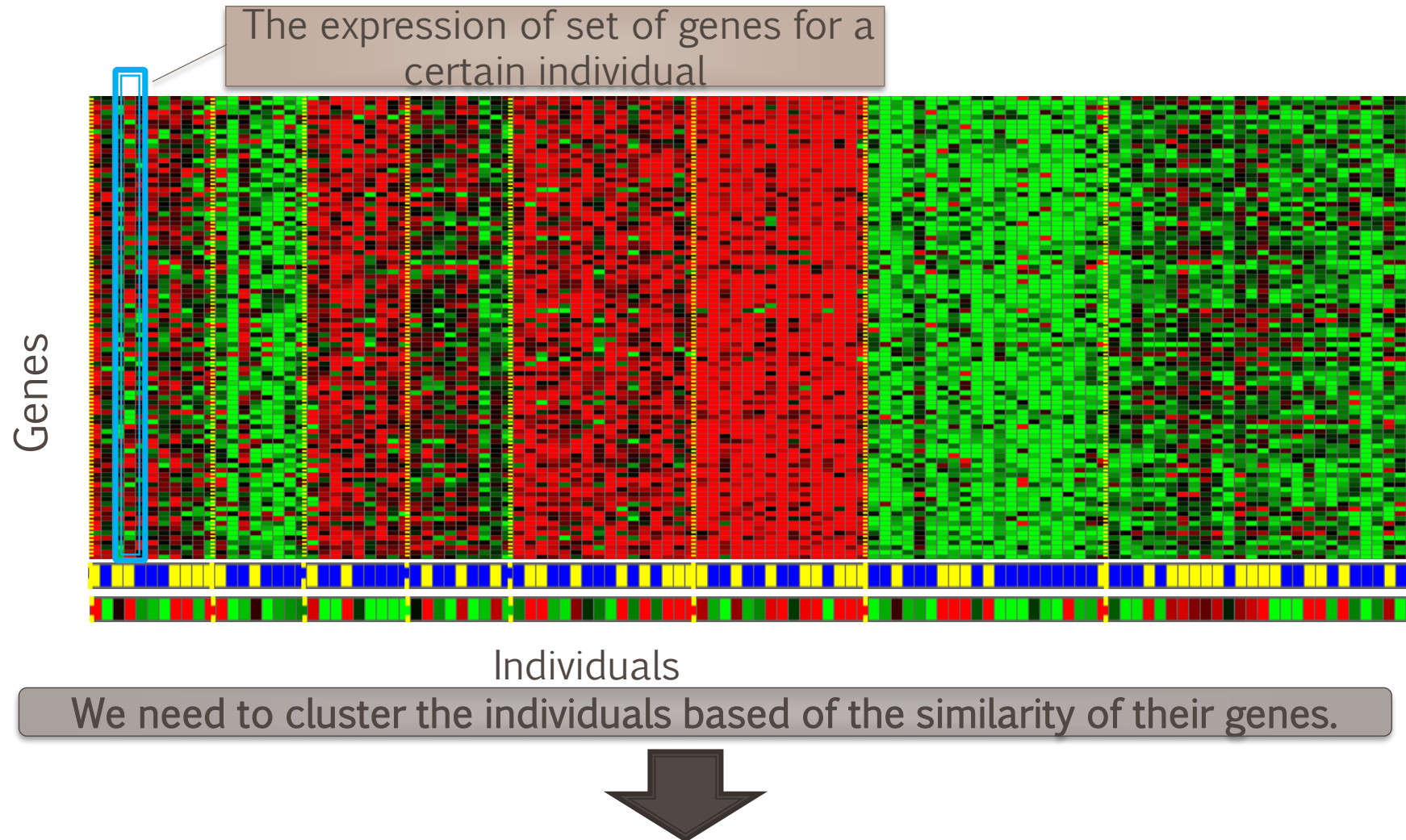


Unsupervised



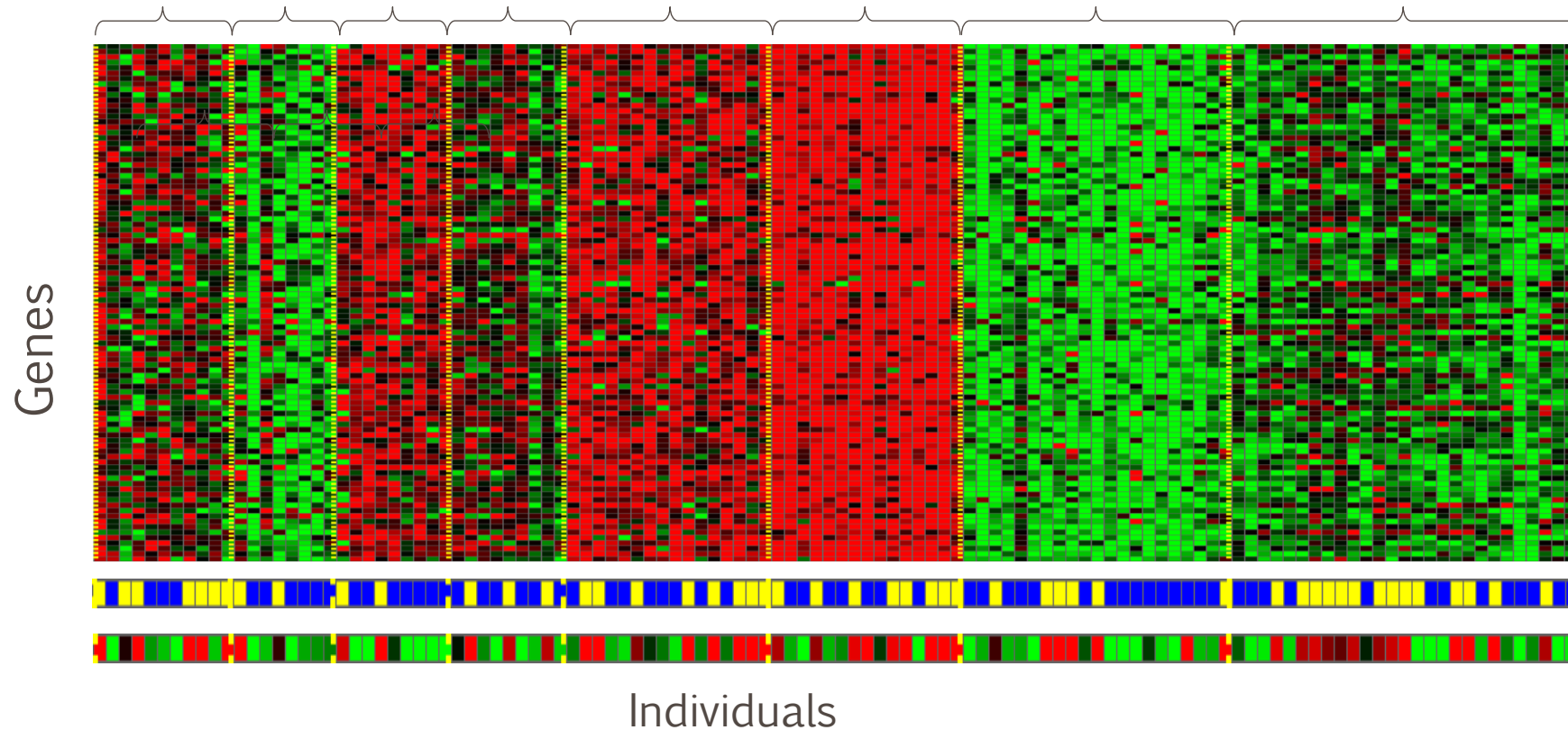
Unsupervised learning is where you only have input data and no corresponding output variables.

Unsupervised Learning applications



[Source: Daphne Koller]

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

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


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 »

 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 YouTube 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 ☆ ▼
CNMMoney - 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

Top Stories (Left Sidebar)

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

Google News

news.google.com

Web Images Videos Maps News Shopping Gmail more

Google news

Search News Search the Web

ADDITIONAL NEWS WEBSITES

U.S. edition Add a section

Top Stories

Deepwater Horizon
Fed meeting
Foreign exchange
market
Lindsay Lohan
IDM
Tom Brady
Tennis
International Film
Festival
Papa Hilton
Jon
Hurricane Igor

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 luring Republicans to the Tea Party. Washington (CNN) - A top White House official sharply denied a report that conservative Obama 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 - 18 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 - AP - CNN - USA Today
all 10 news articles

Deepwater Horizon
BP Oil Well, Site of National Catastrophe, Dies at One
Vanity Fair - Josh Winkler - 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 of Mexico The Associated Press
Wides Doubts BP Would End Operations in Gulf of Mexico Video Dispatching
CNN International - Wall Street Journal (Blog) - The Guardian - New York Times
all 2,262 news articles

Recent

Recession officially ended in June 2009
CNNMoney - Chris Madden - 30 minutes ago
Hurricane Ian lashes Bermuda
USA Today - Gerry Bosman - 5 minutes ago
Explain what you want from us, reads front-page editorial
msnbc.com - Olivia Tones - 10 minutes ago

Crisis response: Pakistan floods

San Francisco Bay Area - E&E

Chavez
Bay Area: Chavez close to sealing STP
San Jose Mercury News - 48 minutes ago
all 24 articles

Google's official search host leaves free company
businessinsider.com
San Jose Mercury News - Chuck Newman - 1 hour ago
Joe Slyke
Martinez man still unconscious as police investigate weekend shooting
San Jose Mercury News - Robert Salgado - 48 minutes ago - all 6 articles

Spotlight

Sarkozy rages at EU humiliation
Financial Times - Peggy Hocking - 8/26/2010

BP Kills Macondo, But Its Legacy Lives On

THE WALL STREET JOURNAL

Log In Register For Free Subscribe Now Get 1 Week Free

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

Article Comments (2)

THE SOURCE HOME PAGE

By James Hiron

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

Fire last response crews battled the blazing remnants of the offshore oil rig last 21, 2010

Photo: Reuters

Most Recent

1. Who Needs Plan II Anyway

2. Will Banks Be Forced to Split Retail And Banking Arms?

3. Timing of Savings Award Incriminating

4. BP Kills Macondo, But Its Legacy Lives On

5. We Already Spent a Small to Small Bill

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

edition.cnn.com

EDITIONS: INTERNATIONAL U.S. FRANCE ARABIC

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

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

By the CNN Wire Staff

September 20, 2010 - Updated 11:17 GMT (2:17 HKT)

Click to play

What next for Gulf oil spill?

STORY HIGHLIGHTS

(CNN) - The ruptured Macondo well, a mile under the Gulf of Mexico, is the largest oil spill in U.S. history, but it's not over yet.

BP oil spill cost hits nearly \$10bn

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 Kabeck
guardian.co.uk, Monday 20 September 2010 08:33 BST
Article history

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

Learning Types

	Supervised	Unsupervised
Discrete	Classification	Clustering
Continuous	Regression	Dimensionality reduction

Machine Learning types

- Semi-supervised learning
 - mix of Supervised and Unsupervised learning
 - usually small part of data is labelled
- Reinforcement learning
 - 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
 - Example - training of self-driving car using feedback from the environment

Grading

Type	topics	grades	date
▪ Assignment 1	Linear & logistic regression	7.5	20/3/2021
▪ Exam 1	Up to Naïve bayes	15	10/4/2021
▪ Assignment 2	Decision tree	7.5	17/4/2021
▪ Assignment 3	SVM	10	1/5/2021
▪ project	NN	20	22/5/2021
▪ Final exam	All contents	40	?
Total	100		