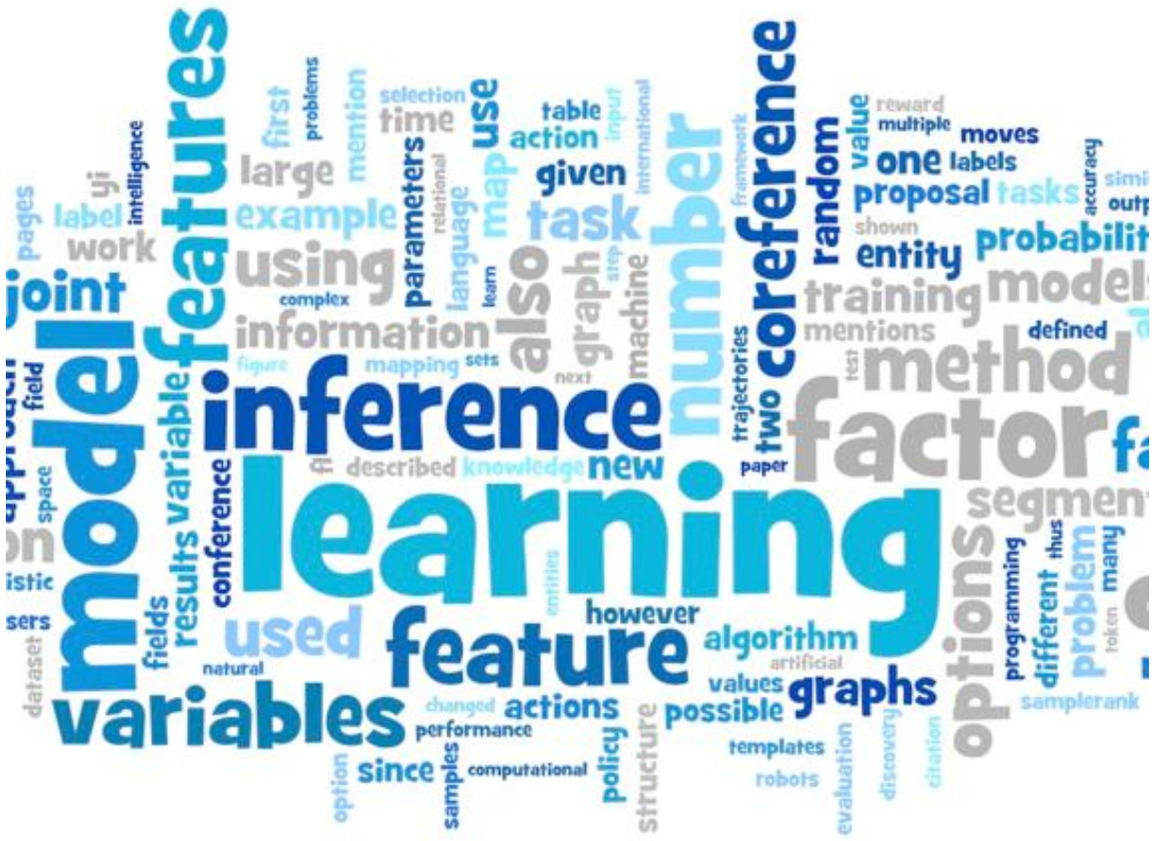
Multimodal Processing, Recognition and Interaction



Summary

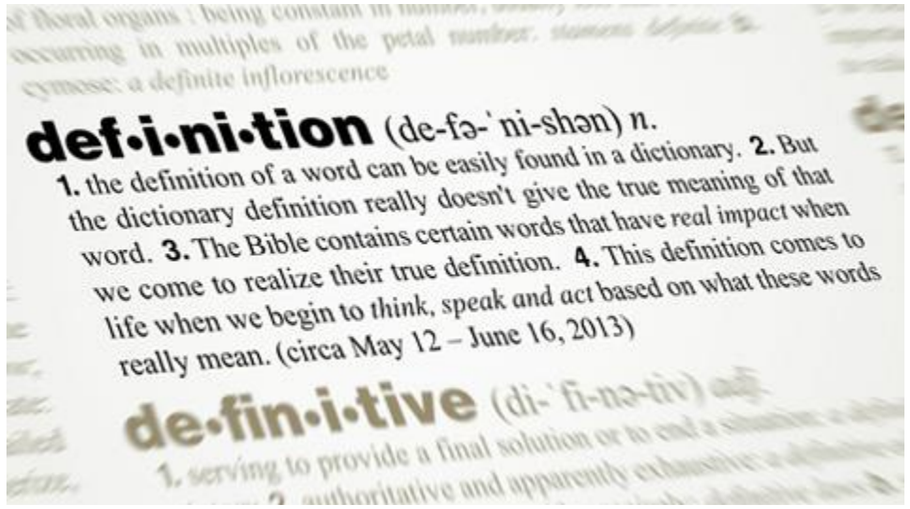
- Multimodal processing, recognition and interaction
- What is machine learning
- Supervised learning
- Unsupervised learning
- Machine learning information flow

Multimodality

- Goals:
  - Integrate computational skills of computers in the real world
  - Extend perceptual and cognitive human capabilities

Multimodality: definition

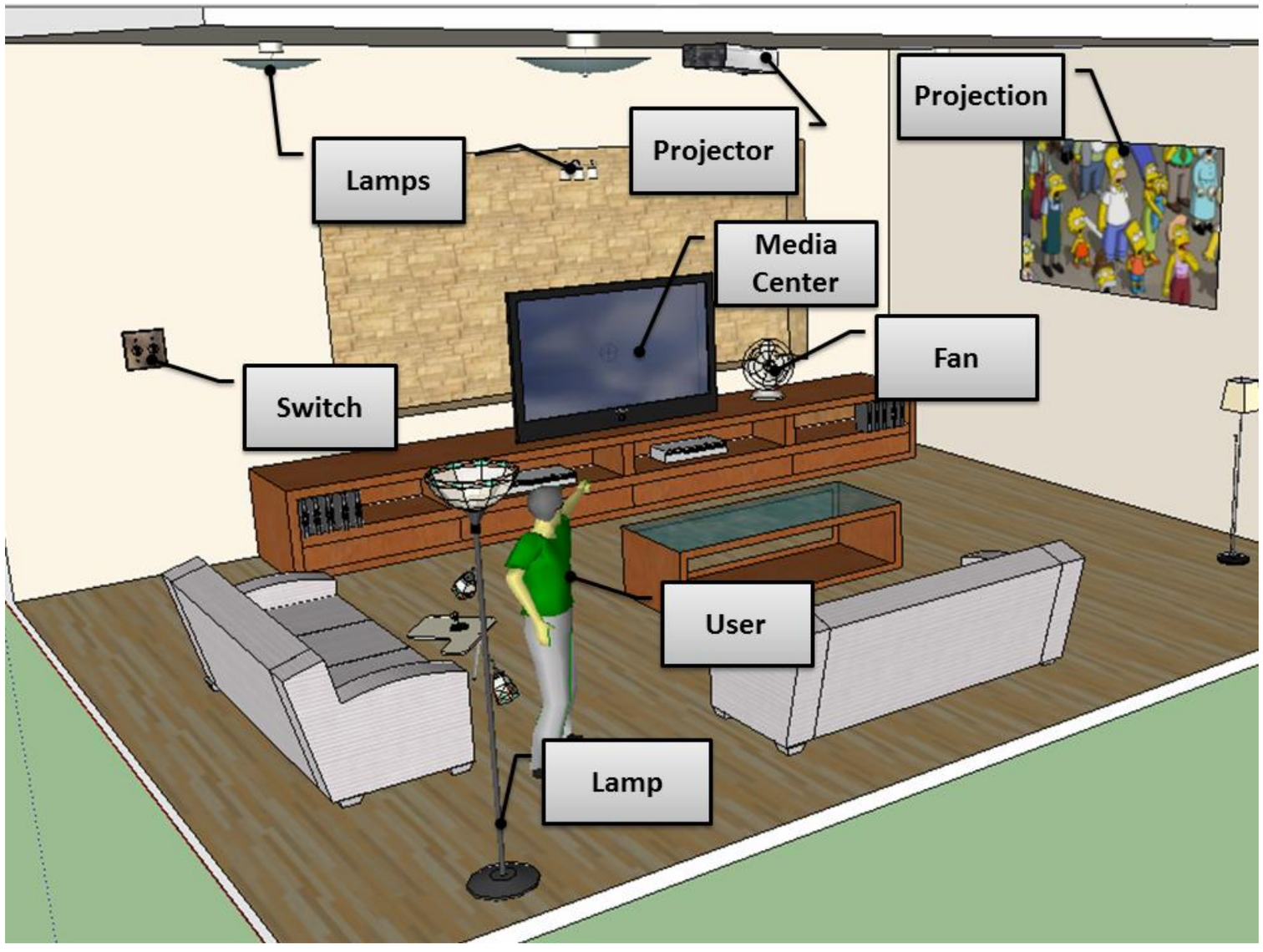
- **Modality** – A channel or path of communication between the human and the computer.

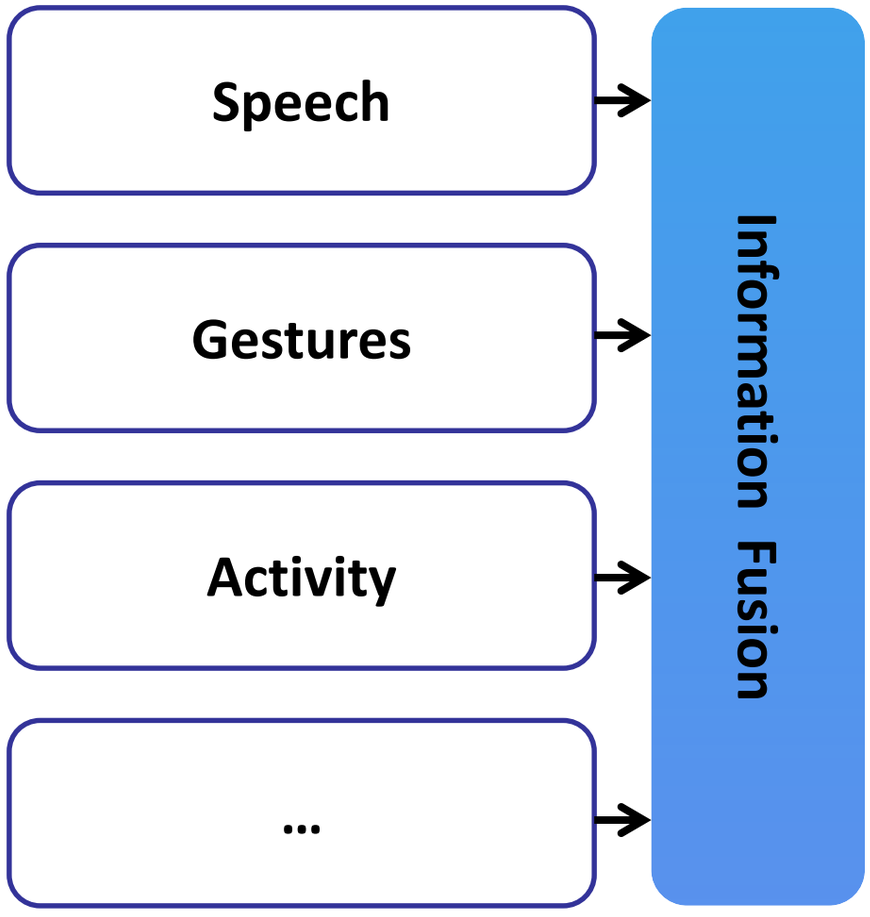


- **Multimodality**[1] – The two features that define a multimodal system are:
  - Fusion of different types of data;
  - Real-time processing: Temporal constraints imposed on information processing;

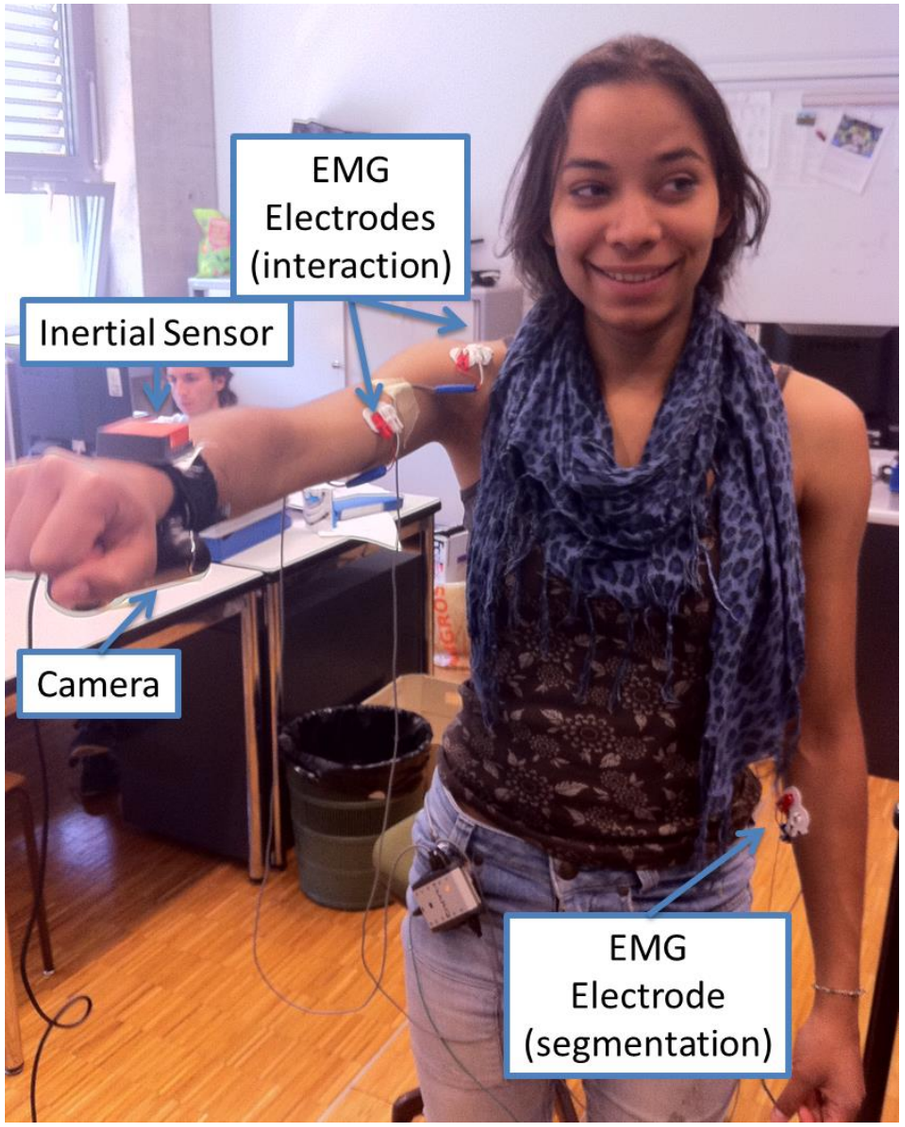
1 - Definition of (Nigay & Coutaz, 1993)

Example 1: Multimodality as multiple interaction channels



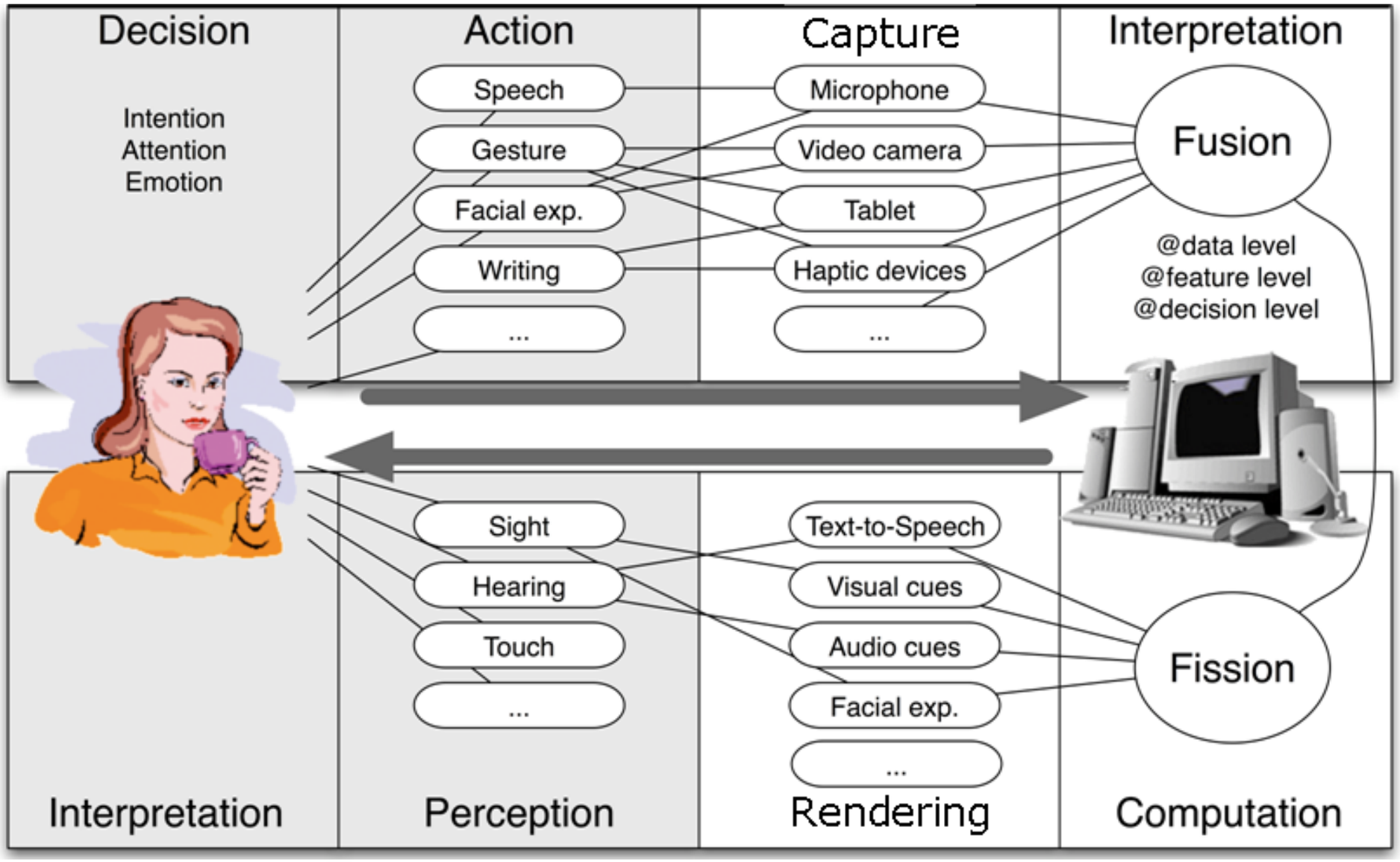


Example 2: Multimodality as fusion of different data types



- Multiple signals:
  - Cameras
  - Accelerometers
  - Physiological data
  - …

Multimodal H/M interaction

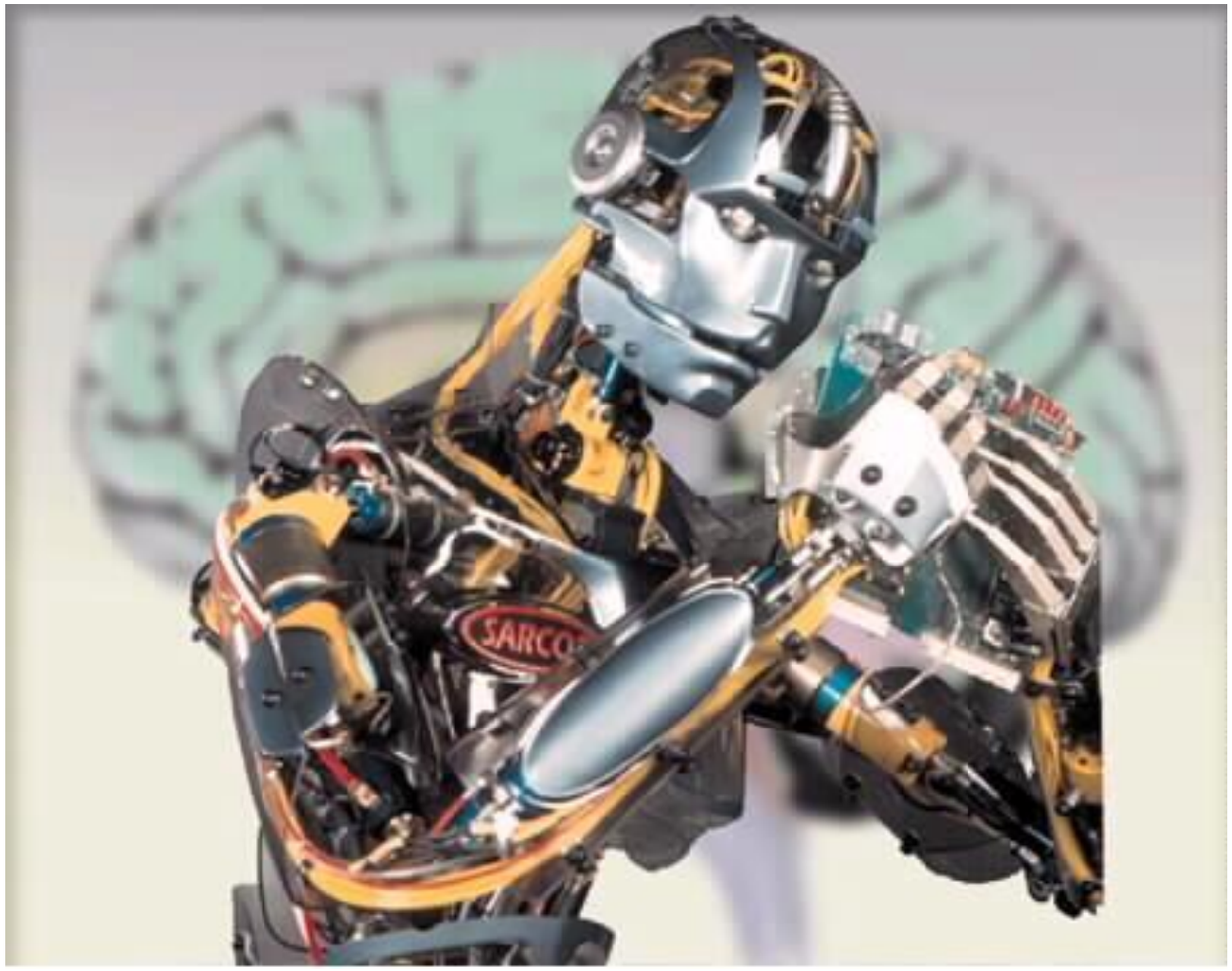


Advantages of Multimodality (potential benefits)

- A list by Maybury and Wahlster [1998, p. 15]:
  - Efficiency
  - Redundancy
  - Perceptability
  - Naturalness
  - Accuracy
  - Synergy

  - Mutual disambiguation of recognition errors [Oviatt, 1999a]

What is machine learning?

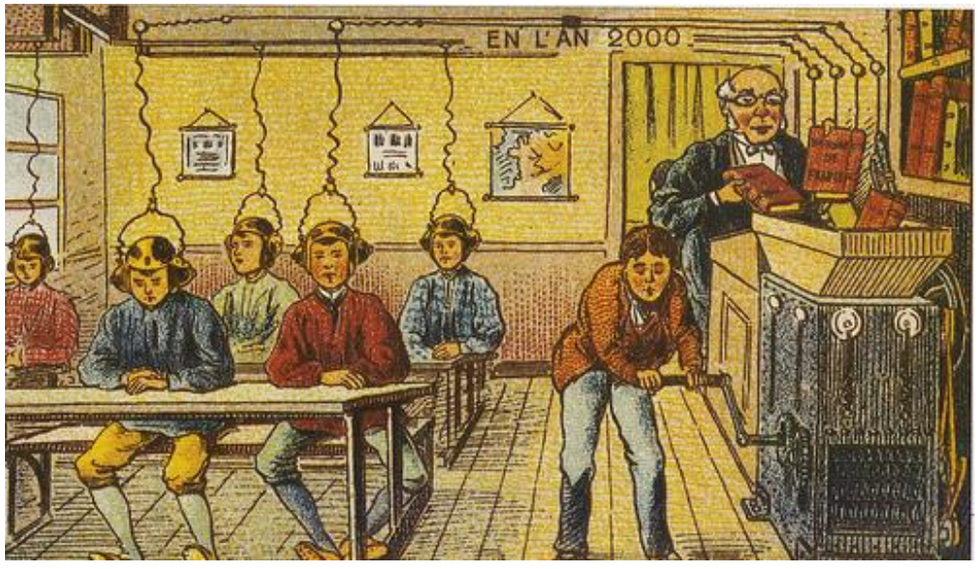


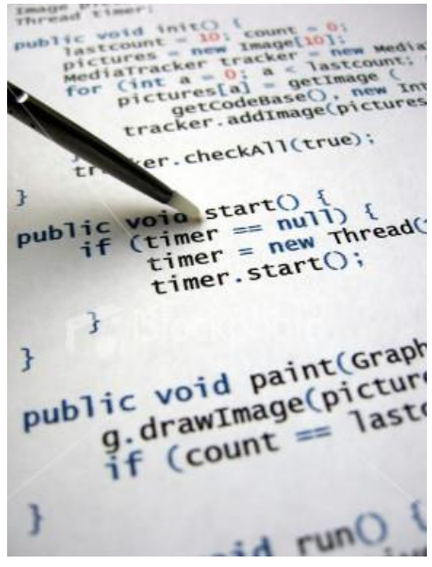
Why Machine Learning

- **3D objects recognition**
  - Different points of view
  - Different illuminations
  - Occlusion
  - Cluttered scene
- **Probability that a credit card transaction is fraudulent**
  - Not rules that are both simple and reliable
  - Need to combine a very large number of weak rules
  - Fraud is a moving target… program needs to keep changing

What is Machine Learning

- The real question is what is ***learning***?
  - *Using past experiences to improve future performance*
- For a machine, experiences come in the form of data
- What does it mean to improve performance?
  - Learning is guided by an objective, associated with a particular notion of loss to be minimized (or, equivalently, gain to be maximized)

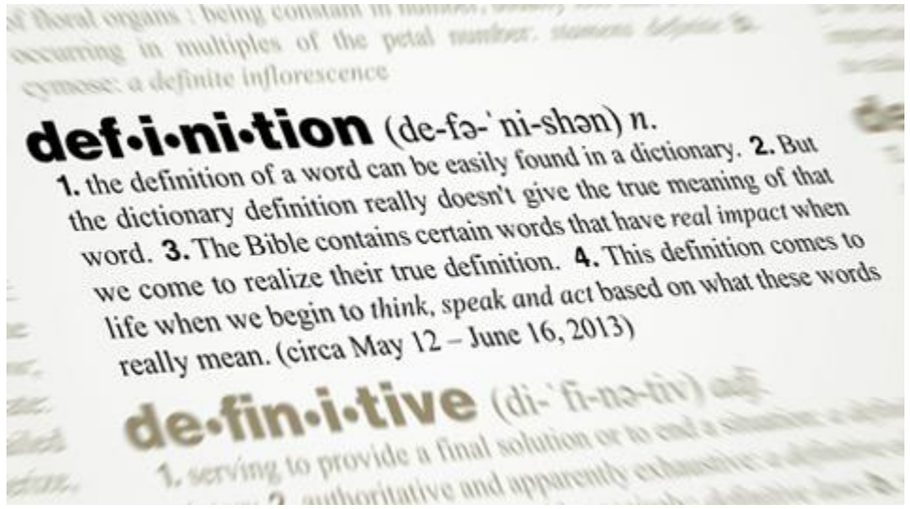


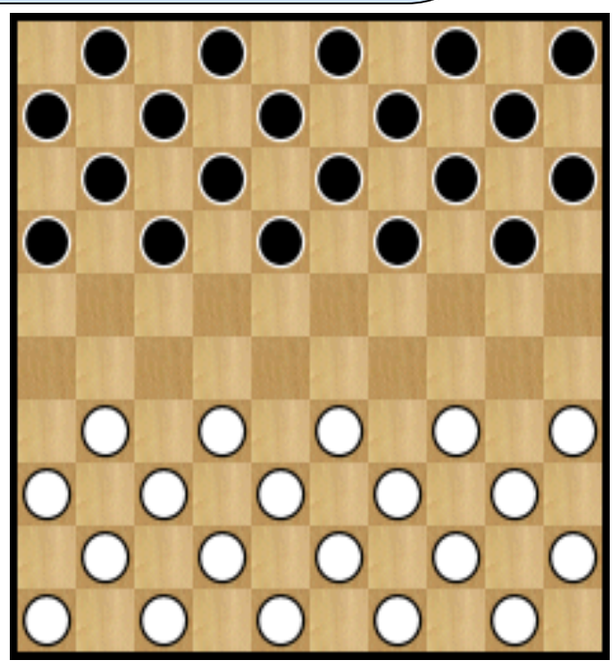


VS

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 (2)

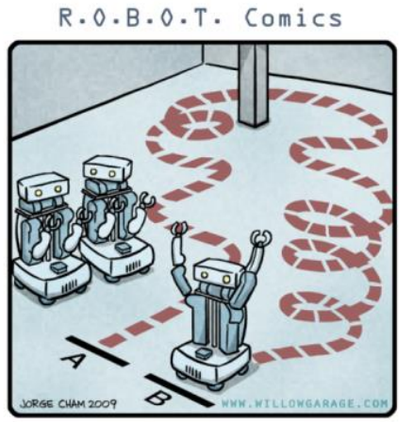
- *Tom Mitchell (1998).*
  - Well-posed Learning Problem: A computer program is said to ***learn*** from experience **E** with respect to some task **T** and some performance measure **P**, if its performance on **T**, as measured by **P**, improves with experience **E**.

Example of tasks (best) resolved by learning

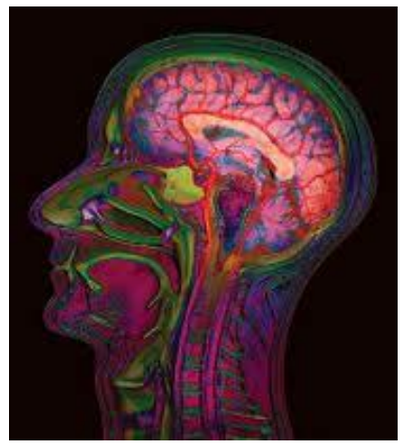
- Recognizing patterns:
  - Objects in real scenes
  - Facial identities or facial expressions
  - Spoken words
- Recognizing anomalies:
  - Unusual sequences of credit card transactions
  - Unusual patterns of sensor readings
- Prediction:
  - Future stock prices or currency exchange rates
  - Which movie will a person like

Applications domains

Recognition



Planning



Diagnosis

Robot Control

Prediction

Typologies

- Machine learning algorithms:
  - **Supervised** learning
  - **Unsupervised** learning
  - Reinforcement learning
- Others:
  - **Semi-supervised** learning
  - Recommender systems.
- We will talk about:
  - Practical advice for applying learning algorithms.

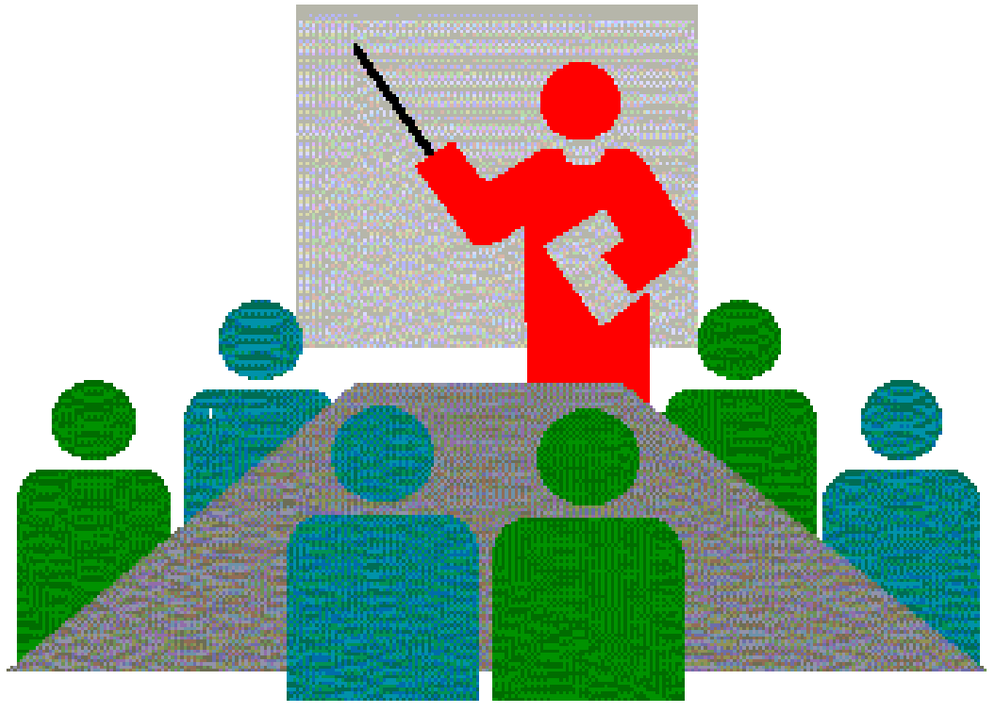
This course

- … will deal with 5 main topics
  - Classification
  - Prediction
  - Anomaly detection
  - Clustering
  - Generative (artistic) approaches

- Real world problems and use cases
- Real world data

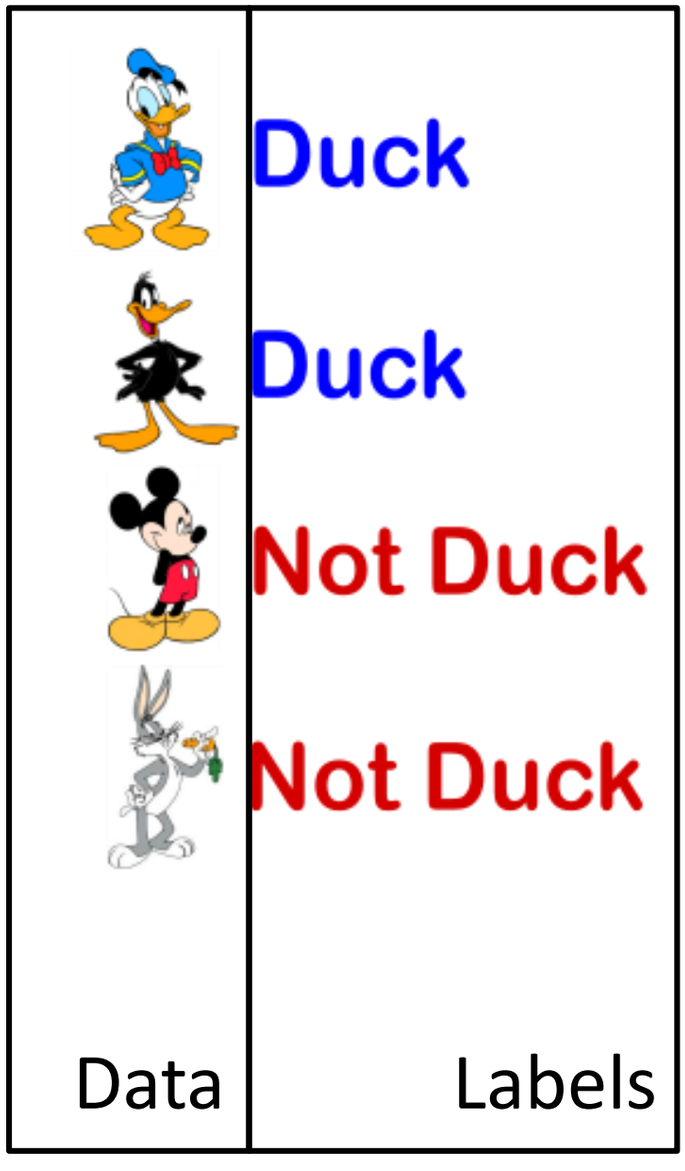
- But before starting basic knowledge is required…

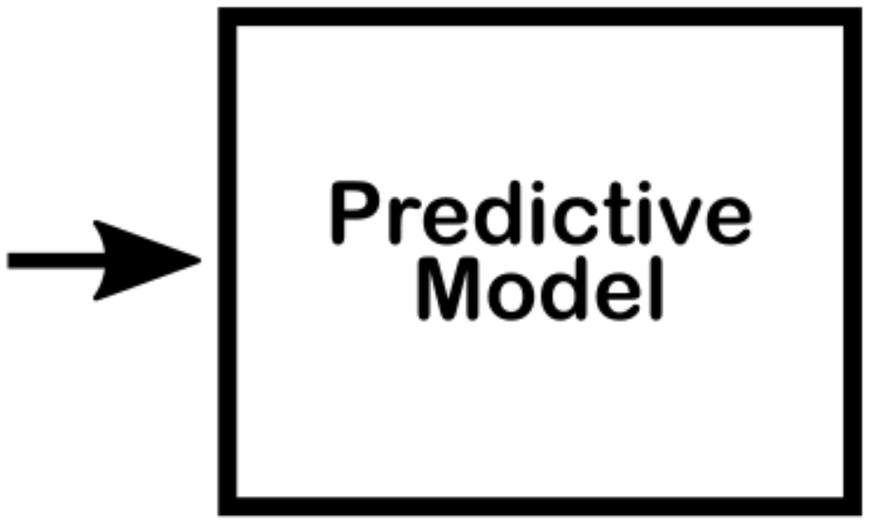
Supervised Learning



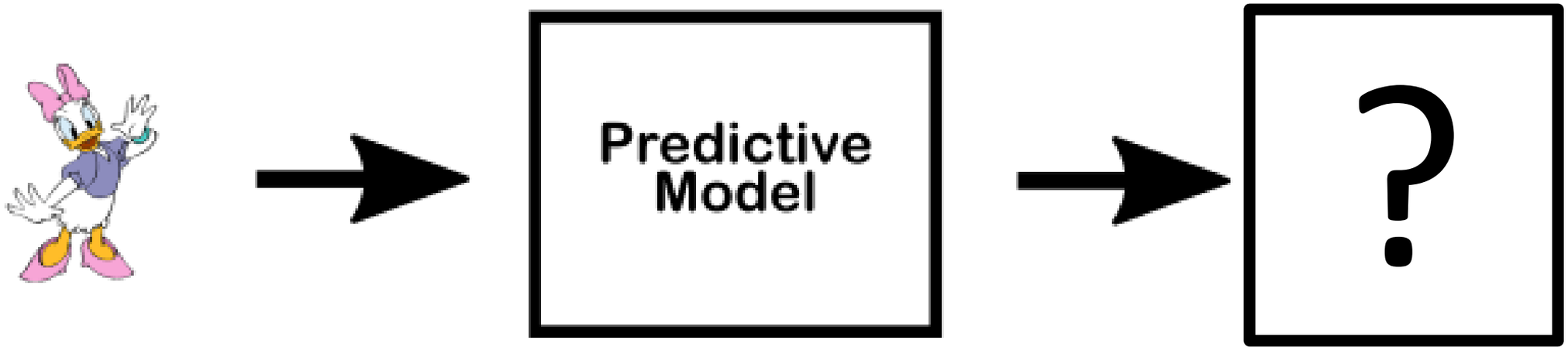
Supervised Learning

Supervised Learning

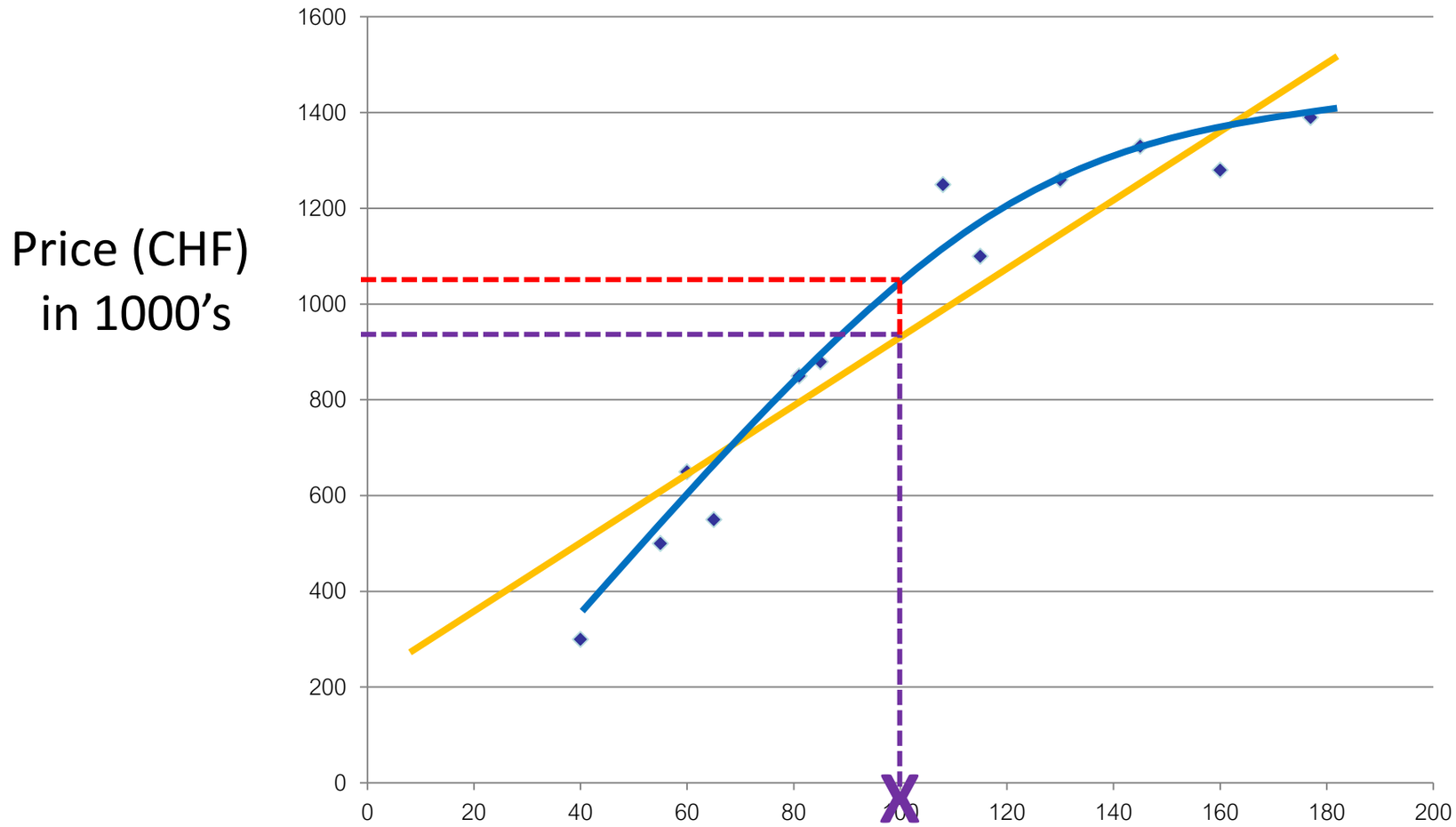




Supervised Learning



Housing price in Lausanne - prediction

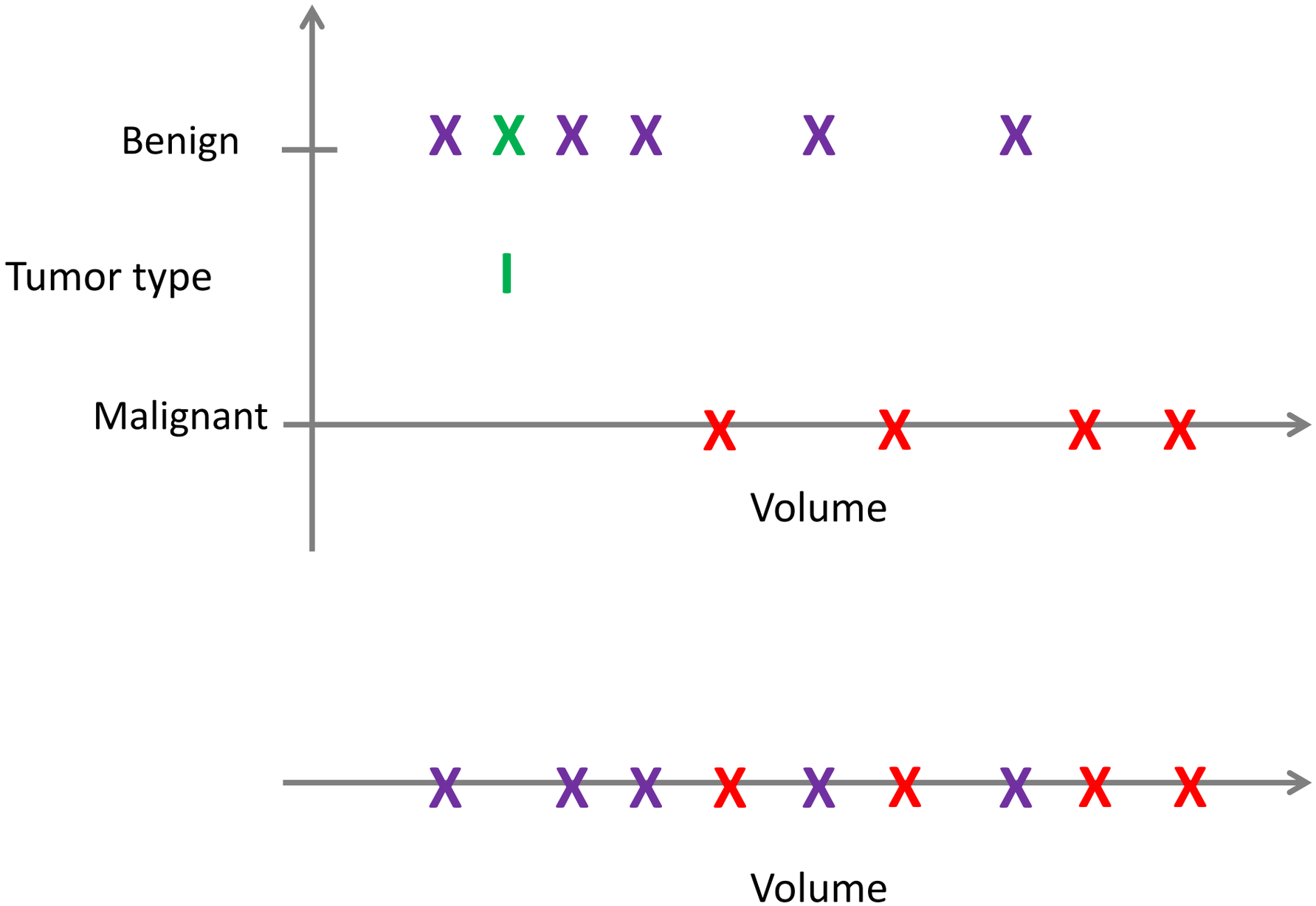


Supervised Learning:

"right answers" given for the training

Regression: Predict continuous valued output (price)

Tumor - Benign, Malignant ?

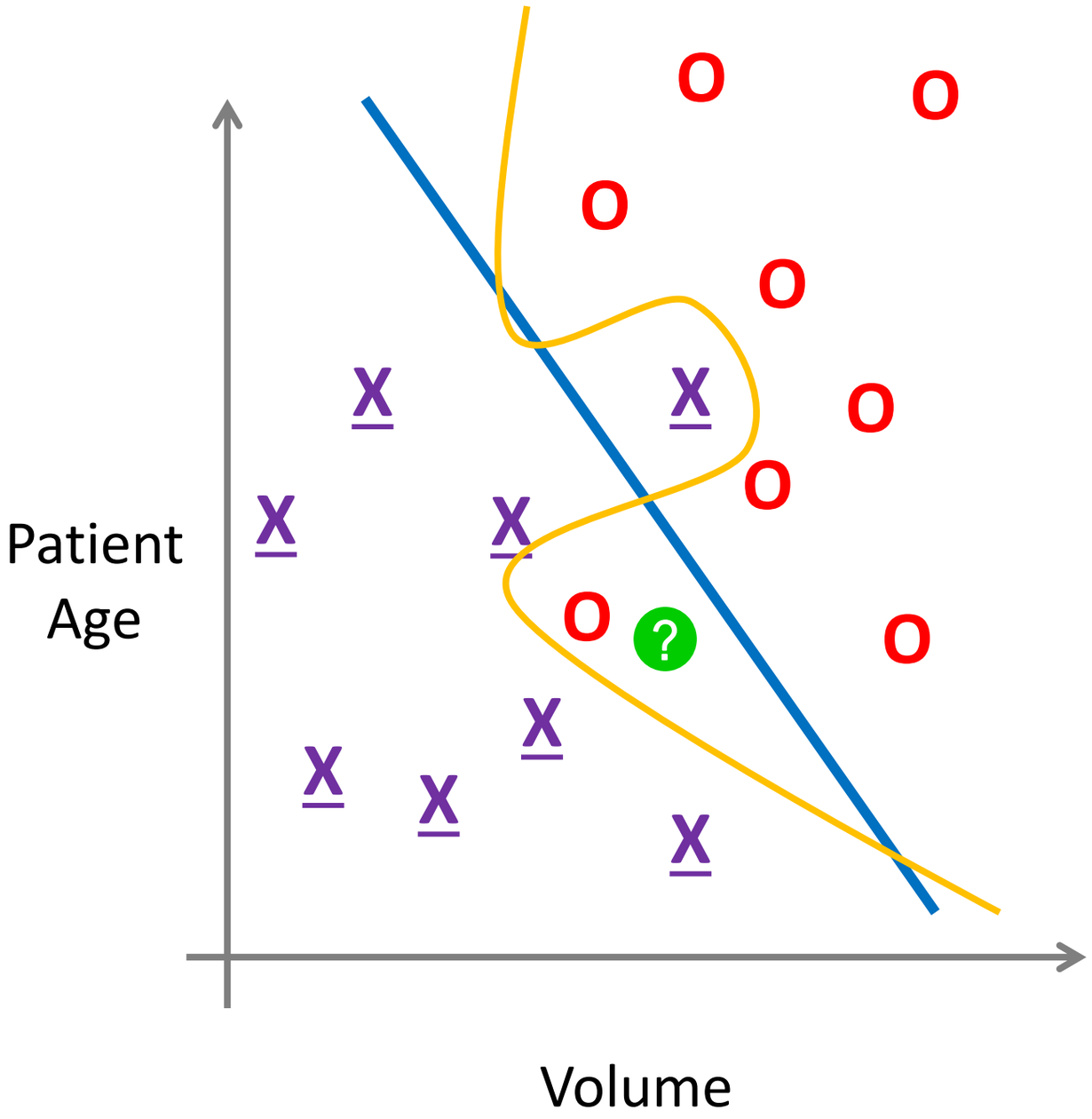


Classification

Discrete valued output

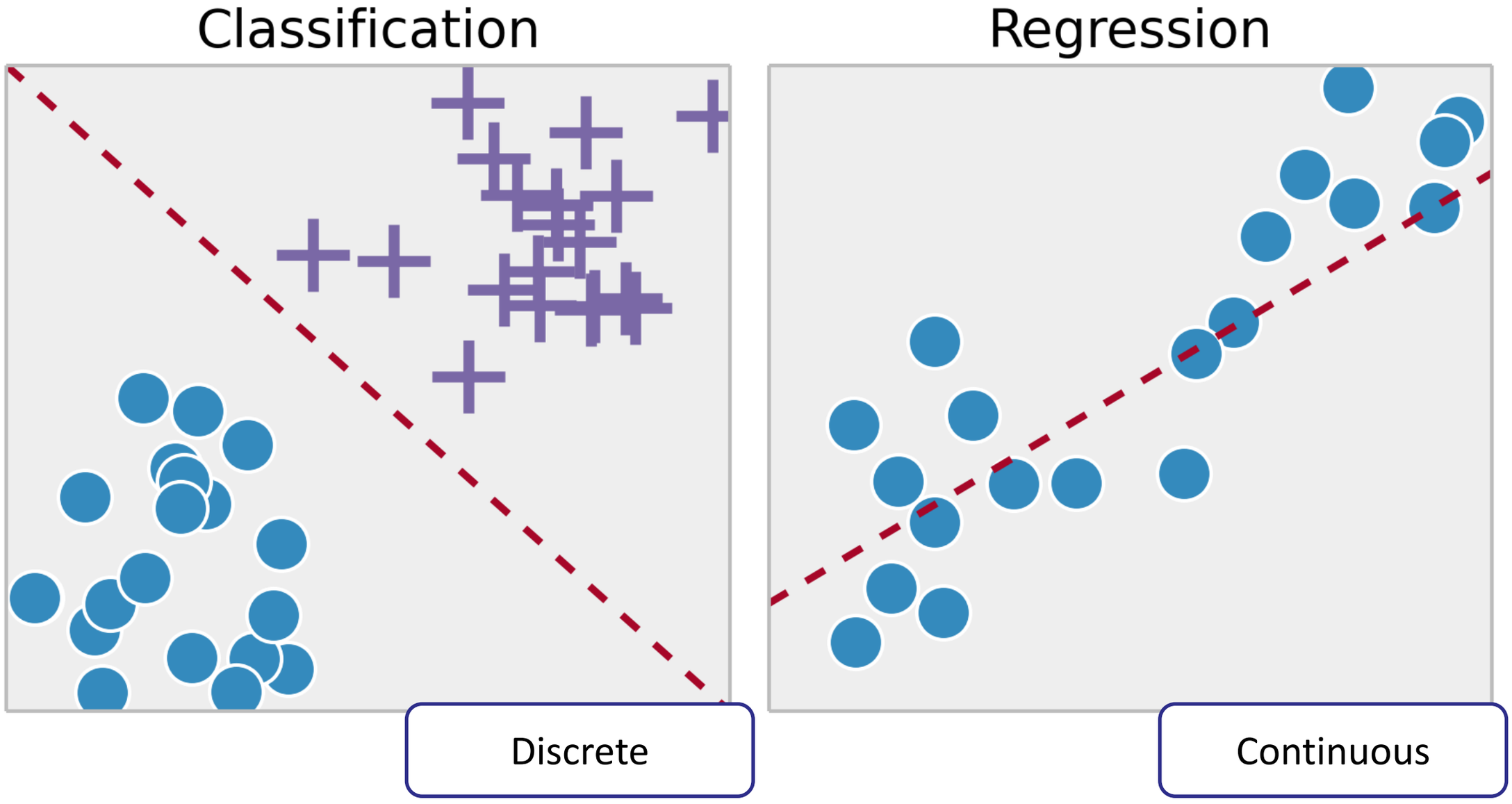
(in this case: B or M)

Tumor (Benign, Malignant)



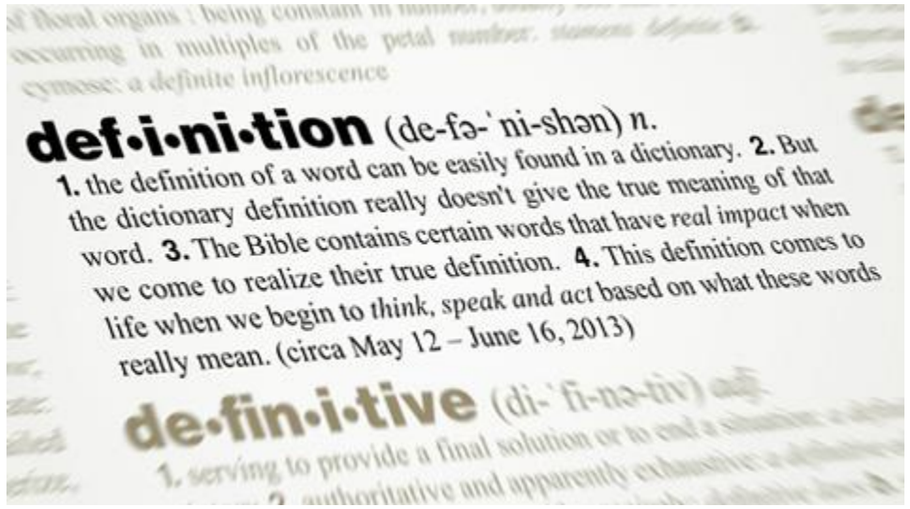
- Examples of **features**
  - Smoking habits
  - Family history
  - …

To summarize



Supervised learning: definition

Part of the data (observations, measurements, etc.) are labeled with pre-defined classes/values.



It is like that a “teacher” gives the classes (supervision).

Question: Regression Vs Classification

You're running a company developing learning algorithms. You should address of three problems for your customers:

Problem 1: Company_1 has a database with information about the sells of the last 3 years. You want to predict how many products Company_1 will sell over the next 3 months.

Problem 2: Company_1 comes back to you with the same database with information about the sells of the last 3 years. You want to predict if Company_1 will sell FEW, AVERAGE or MANY products over the next 3 months.

Problem 3: Company_2 needs a software to analyze employees e-mail accounts, and for each account filter out spam messages.

Classes and features

- Classes:
  - Def: *in a classification problem a* **class** *is the label that we want to recognize*

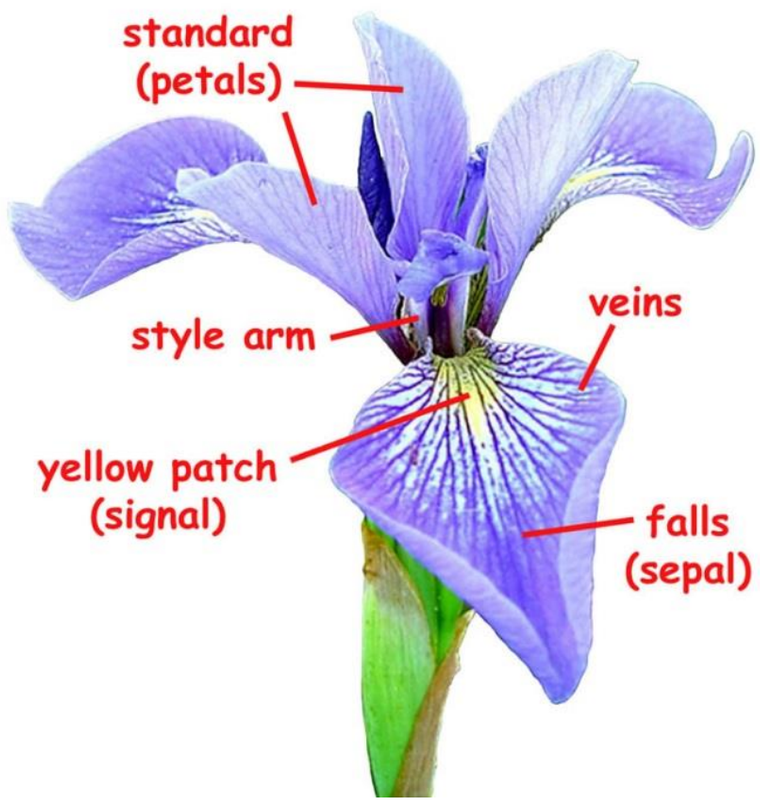
Classes and features

- Classes:
  - Def: *in a classification problem a* **class** *is the label that we want to recognize*



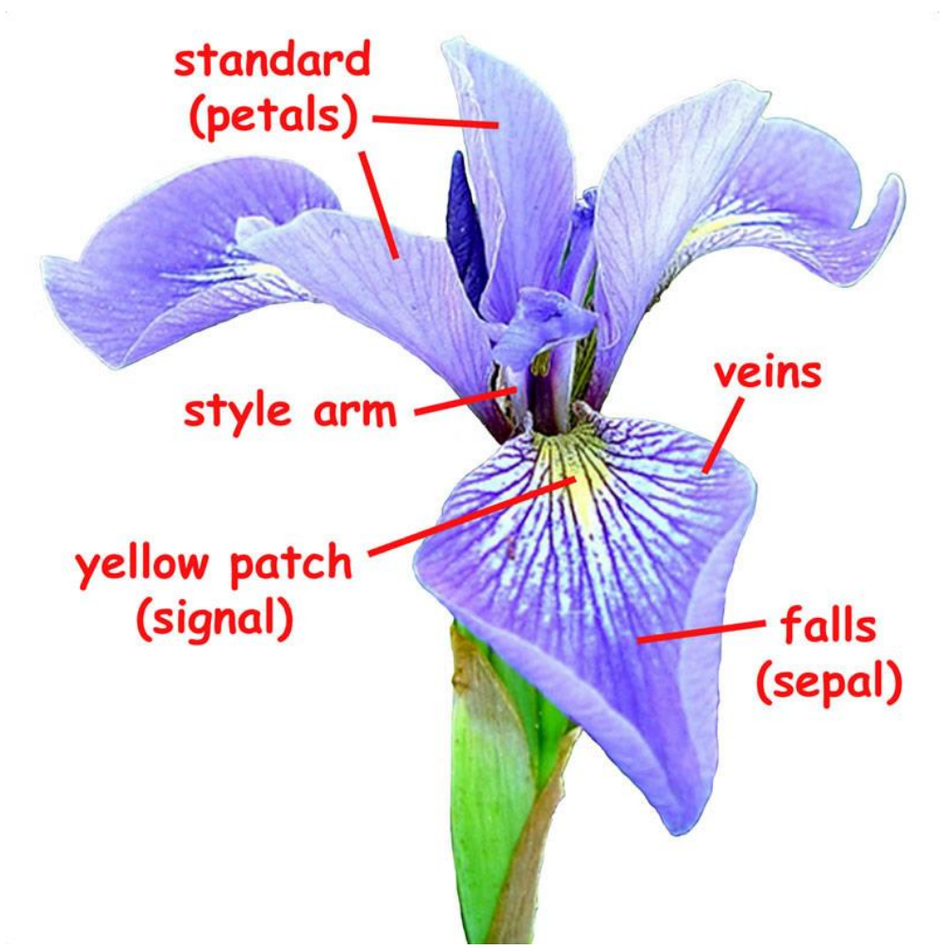
Classes and features

- Features:
  - Def: *A* feature *is the specification of an attribute and its value.*
    - For example, *color* is an attribute. ``Color is blue'' is a feature of an example.



Classes and features

➢ Number of features?

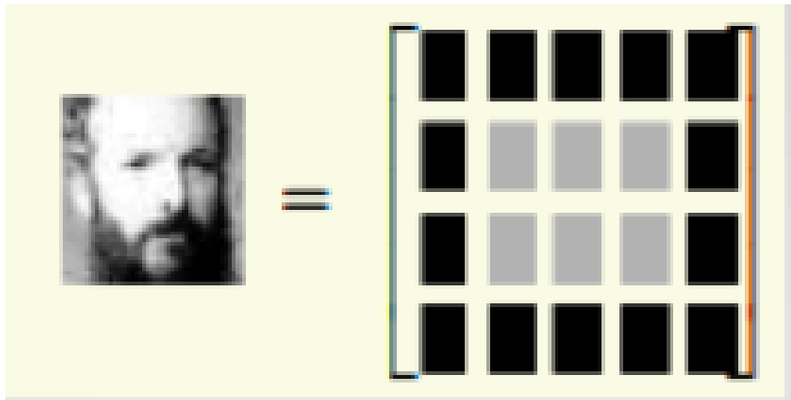


- Color
- Size
- Number of pistils
- Veins colors
- Stem color
- Stem width
- …

In a computer vision application:
Simply all the pixels in the image!

Or not?

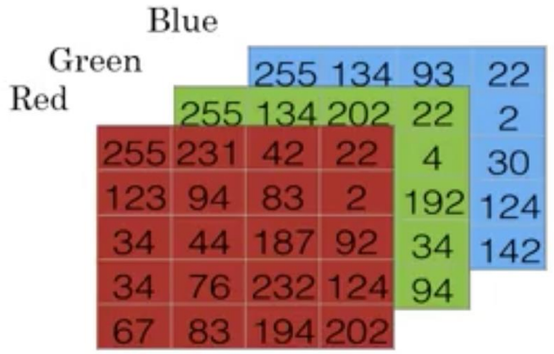
Classes and features



Example: Face detection

Classes: faces/no face

Features: image pixels



Ex. 20 by 20 pixels = 400 features!

Are these 400 features really interesting??

=> Only the most relevant features should be used!! *

** Or you should have A LOT of data!*

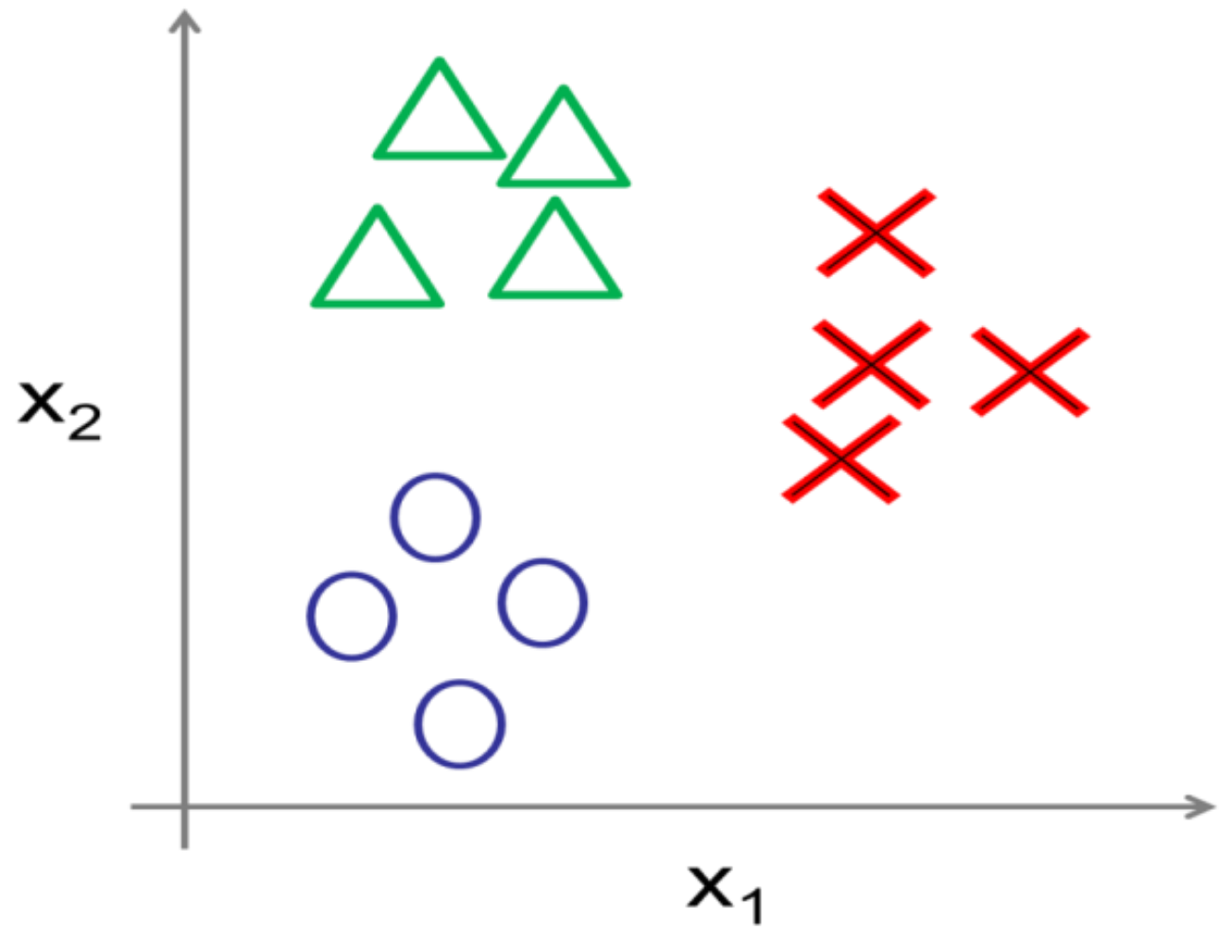
Feature Extraction

- **Essential** step in pattern recognition and machine learning problems.

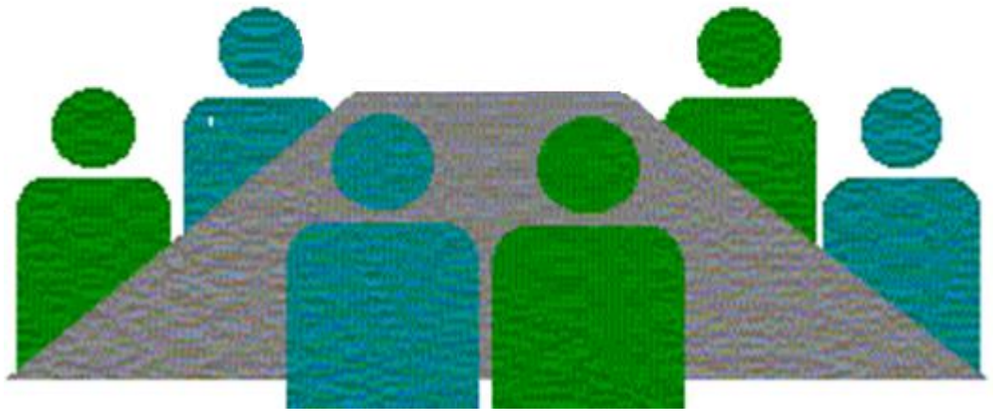
- It is often decomposed into **feature construction** and **feature selection**
  - Feature Construction:
    - Goal: balance expressiveness of features with the size of the corresponding feature space
  - Feature Selection:
    - Goal: Select relevant and informative features

Question: Terminology - Features Vs. Classes

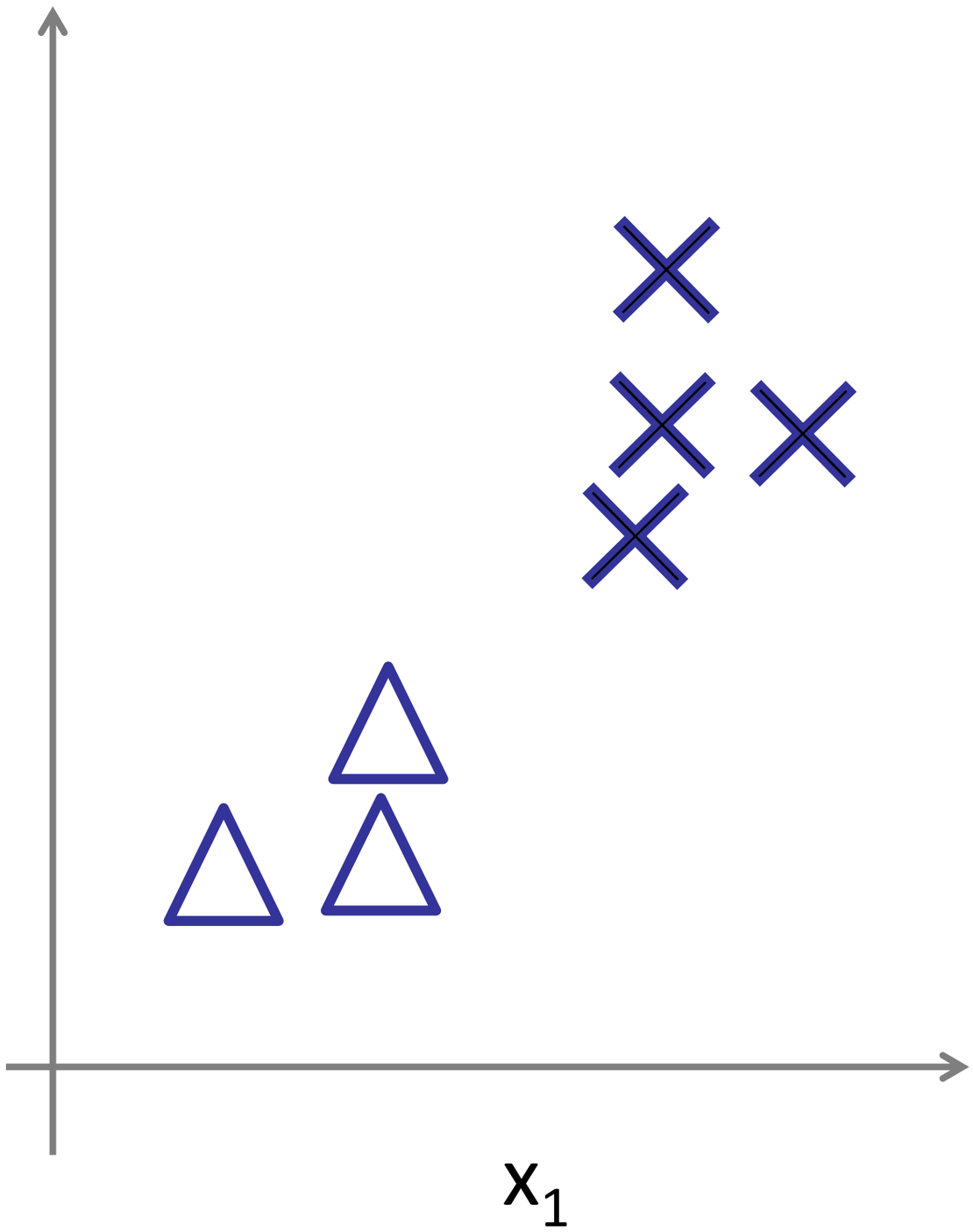
In the following representation, how many classes, how many features?



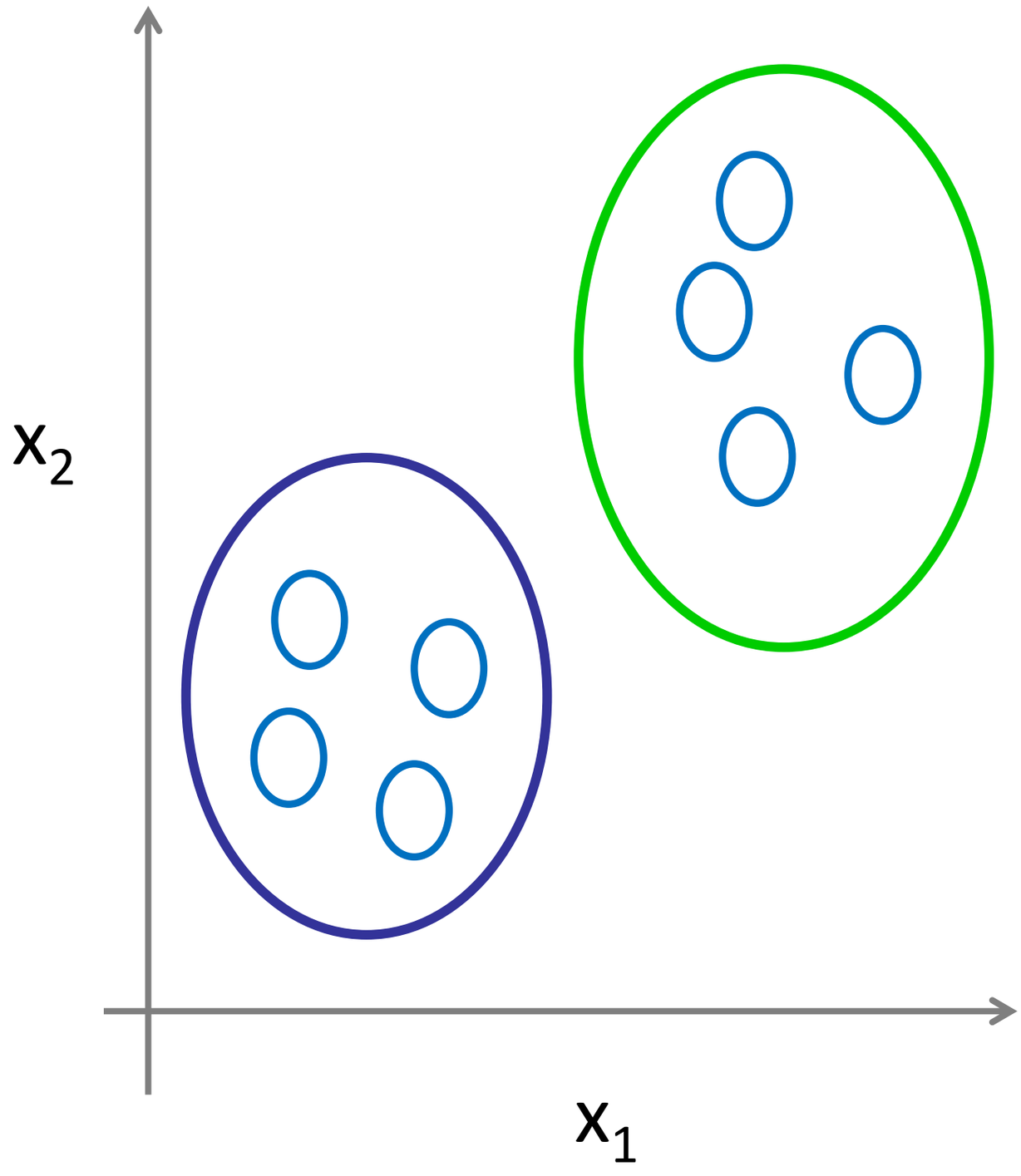
Unsupervised Learning



Supervised Learning



Unsupervised Learning



Unsupervised learning

- Goal: the basic task of unsupervised learning is to **find hidden structure in unlabeled data**

- Why?
  - Labeling is an expensive (sometime impossible) task!

Google News

Search

Headlines | Local | For You | U.S.

S

Top Stories

World

U.S.

Business

Technology

Entertainment

The Emmy Awards

Sports

Science

Health

Manage sections

Top Stories

Three storms raging in Atlantic -- with one aiming for Irma's path

CNN · 5h ago

MORE ABOUT

National Hurricane Center

Hurricane Jose

Hurricane Jose lurks off the East Coast, Tropical Storm Maria threatens the Caribbean

Washington Post · 18h ago

The Latest: Jose moving slowly, causing strong rip currents

In Depth · ABC News · 19h ago

View full coverage →

British Police Make 'Significant Arrest' in Subway Bombing

New York Times · Sep 16, 2017

RELATED COVERAGE

London Tube terror attack: Everything we know so far about Parsons Green bomb

Local Source · Evening Standard · Sep 15, 2017

Here's why a judge acquitted a St. Louis cop of first-degree murder

More Local Ne

Google News has
local news with C
Check out the "Lo

More Local News

In the News

Donald Trump

North Korea

London Undergr

Democratic Part

Republican Part

Japan

Chuck Schumer

Nancy Pelosi

Jason Stockley

Top Stories

MORE ABOUT

National Hurricane Center

Hurricane Jose

Three storms raging in Atlantic -- with one aiming for Irma's path

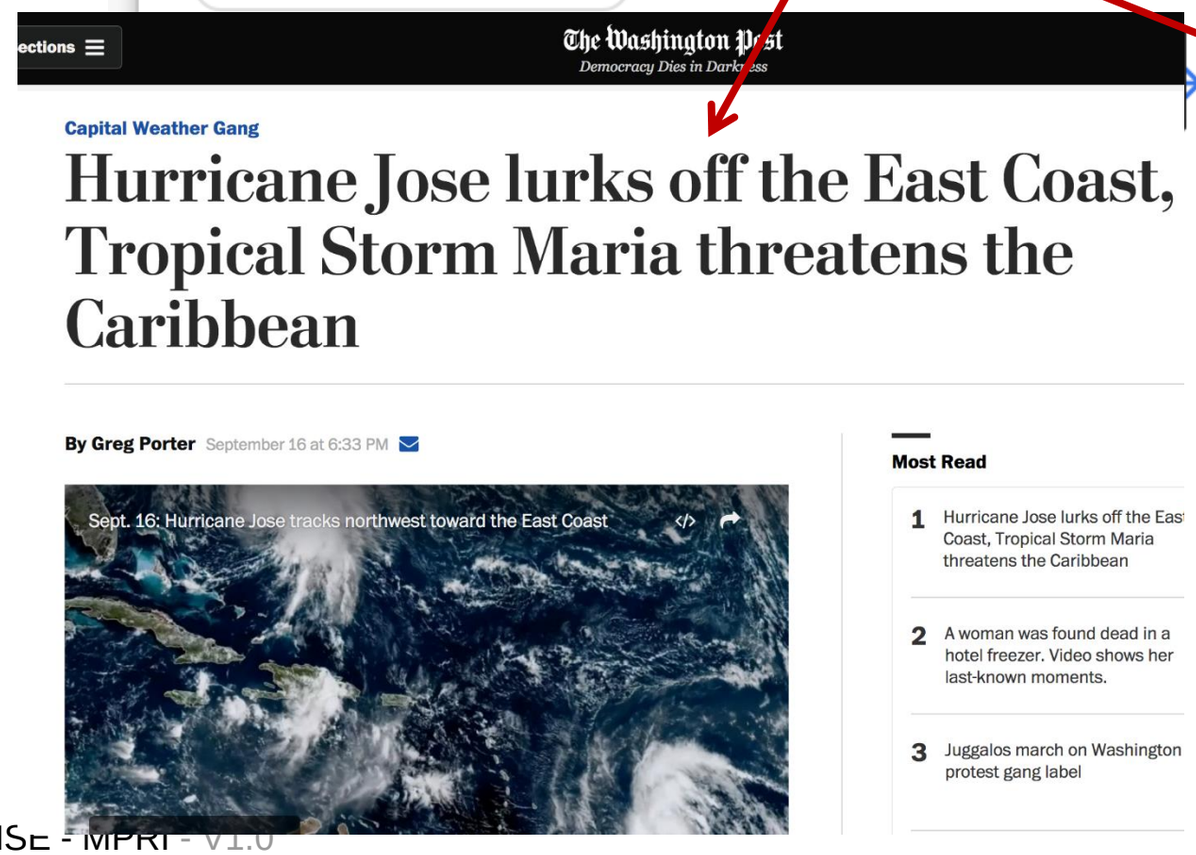
CNN · 5h ago

Hurricane Jose lurks off the East Coast, Tropical Storm Maria threatens the Caribbean

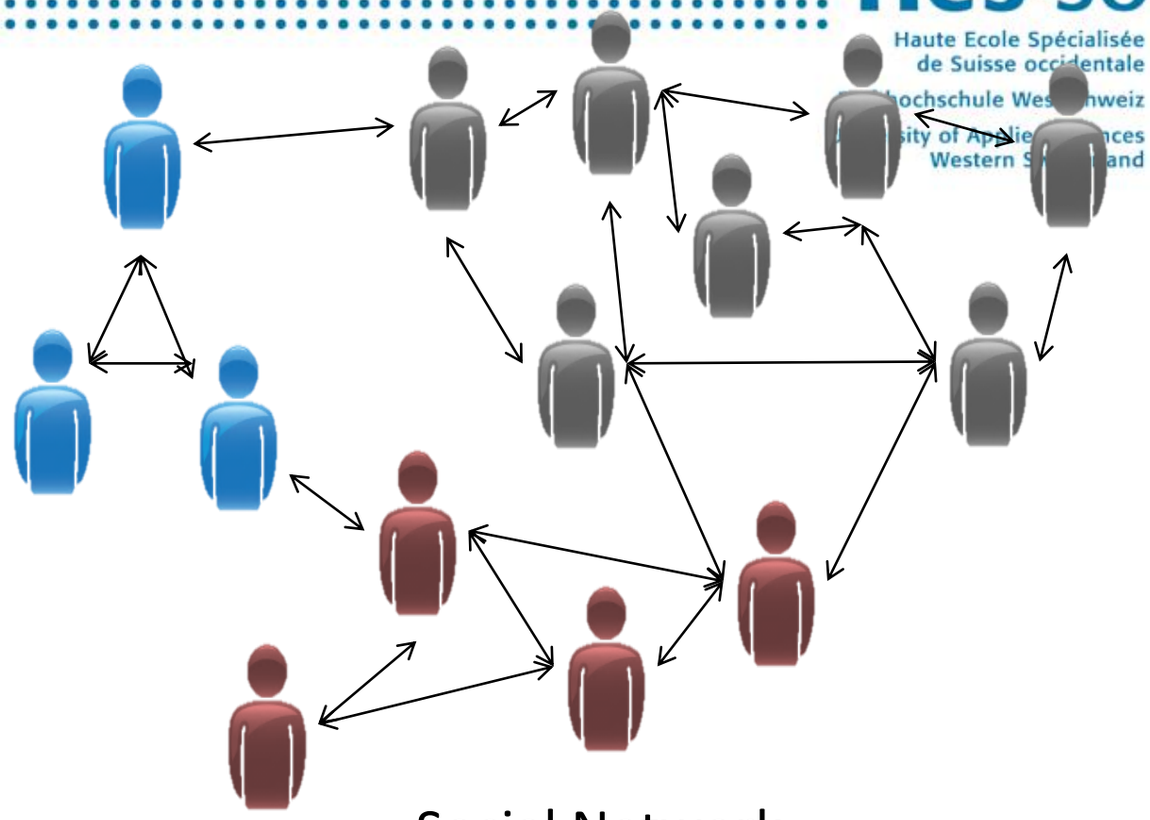
Washington Post · 18h ago

The Latest: Jose moving slowly, causing strong rip currents

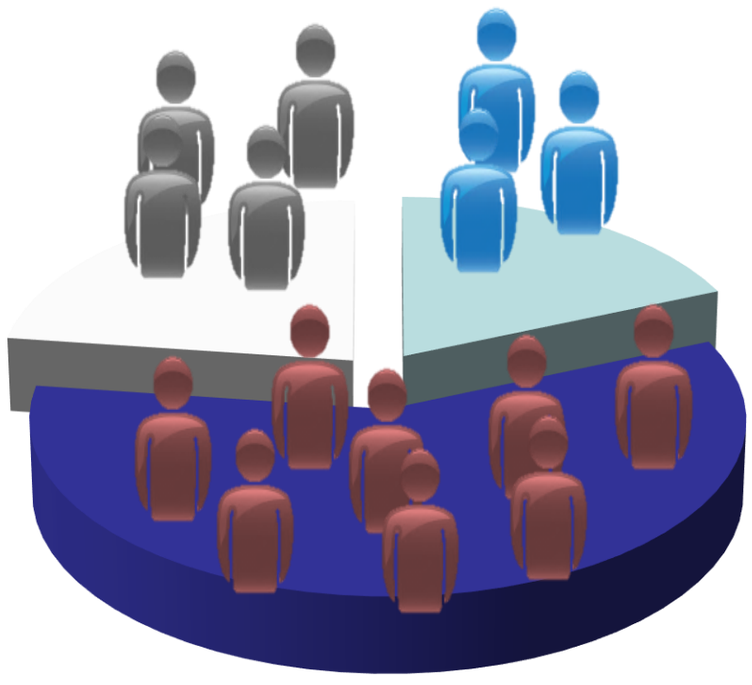
In Depth · ABC News · 19h ago



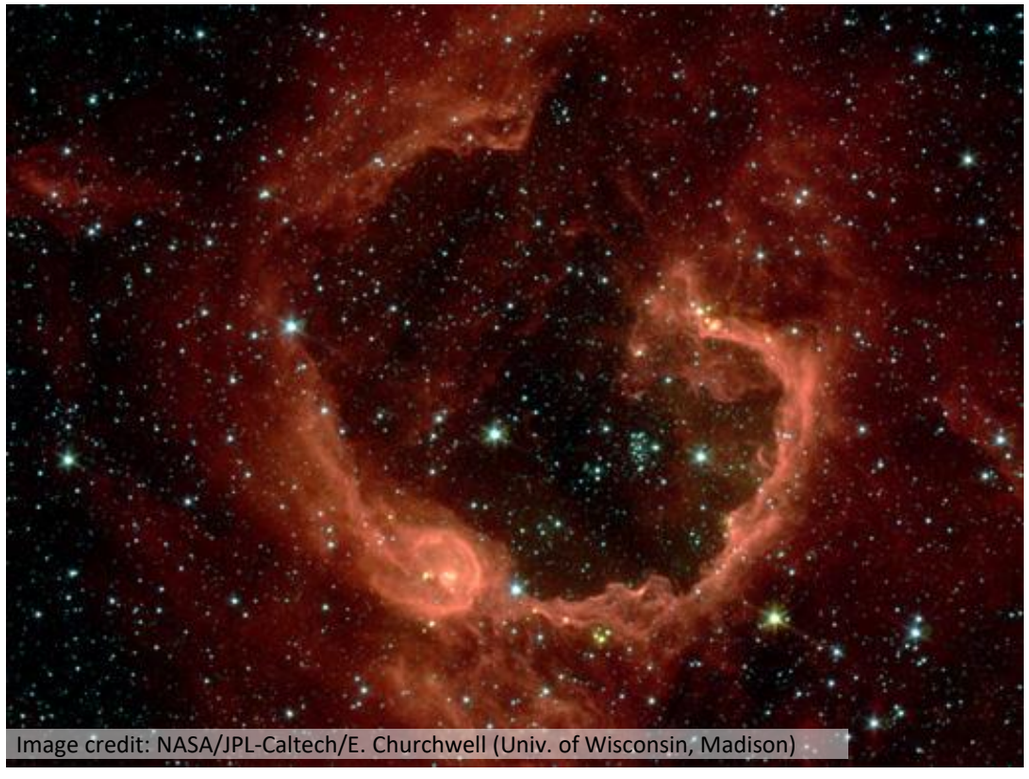
Organize computing clusters



Social Network



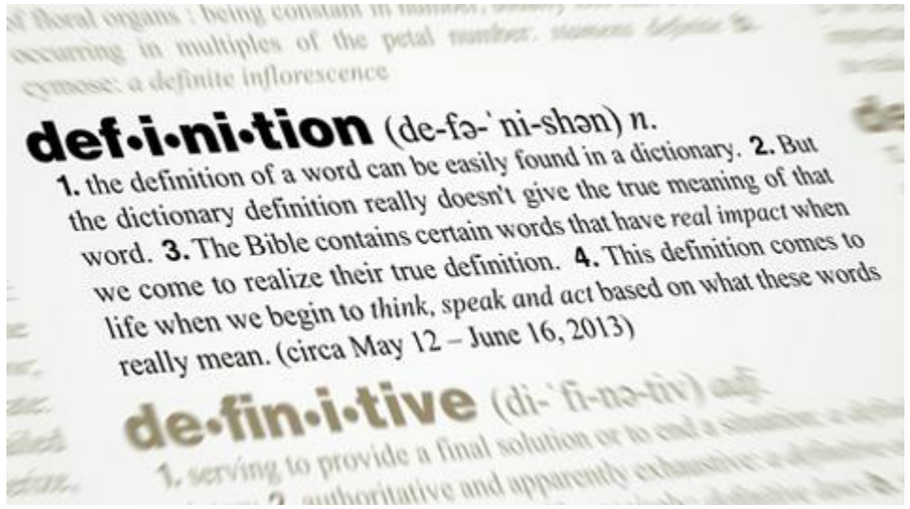
Market segmentation



Astronomical data analysis

Unsupervised learning: definition

- The term “unsupervised learning” or “learning without a teacher” is generically associated with the idea of using a collection of observation $X_1,\dots X_n$ sampled from a distribution p(X) to describe properties of p(X).



- Definition of: **Vittorio Castelli http://www.ee.columbia.edu/~vittorio/**

Of the following examples, which would you address using an unsupervised learning algorithm?

- **Given email labeled as spam/not spam, learn a spam filter.**

- **Given a dataset of not labeled images cluster similar pictures.**

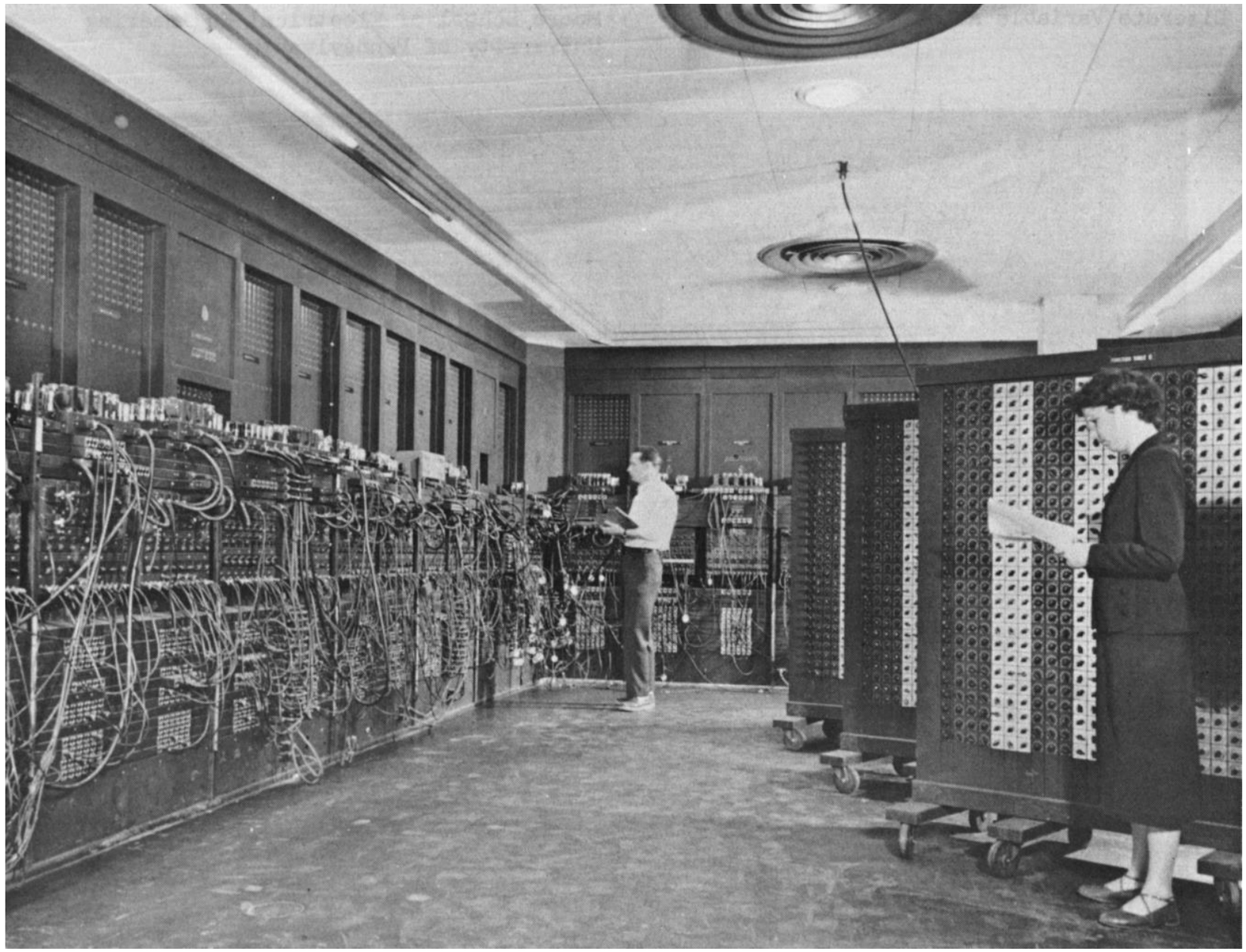
- **Given a database of customer data, automatically discover market segments and group customers into different market segments.**

- **Given a dataset of patients diagnosed as either having diabetes or not, learn to classify new patients as having diabetes or not.**

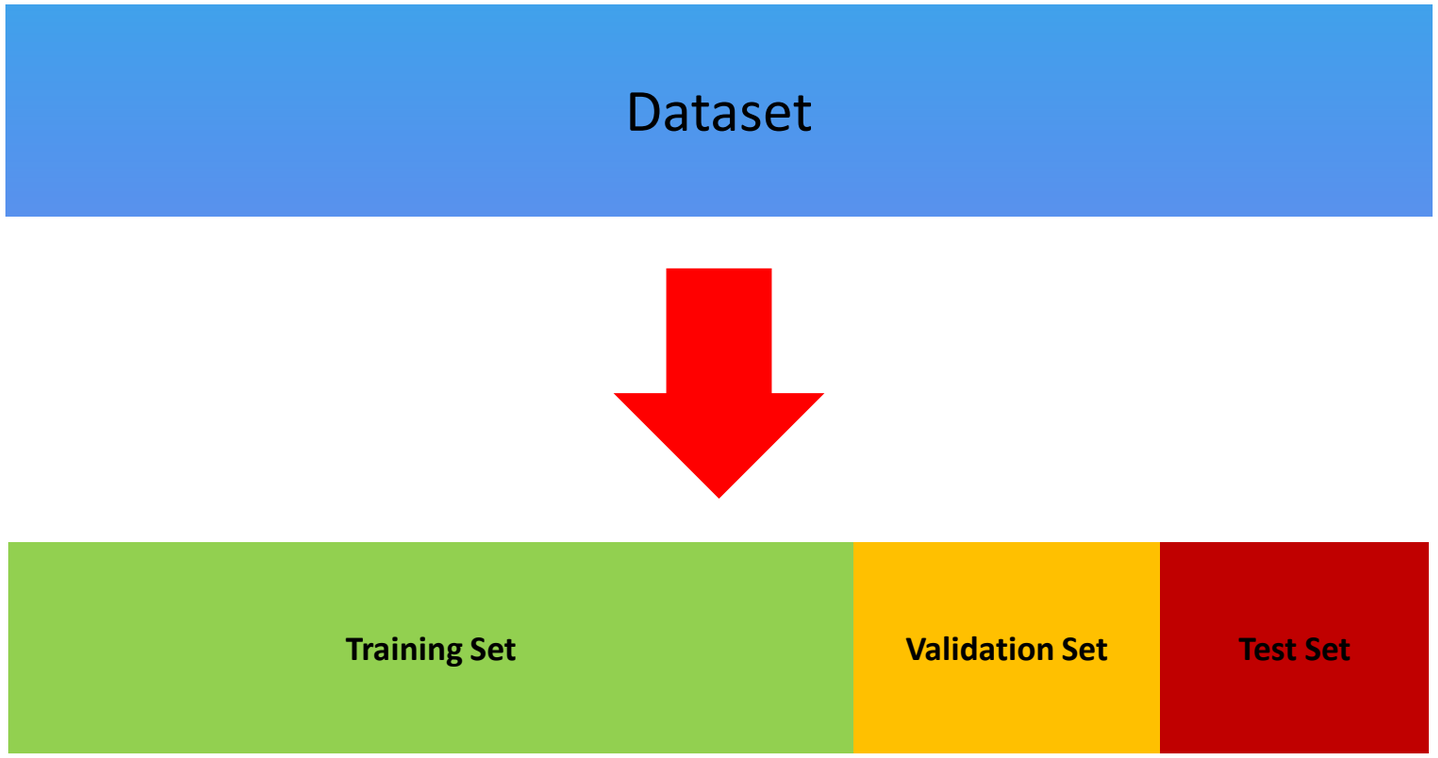
Summary:

- Supervised learning
  - Learn to produce an output when given an input vector
  - Two flavors:
    - Regression: the target output is a real number or a whole vector of real numbers
    - Classification: the target output is a class label
- Unsupervised learning
  - Discover a good internal representation of the input

Information processing



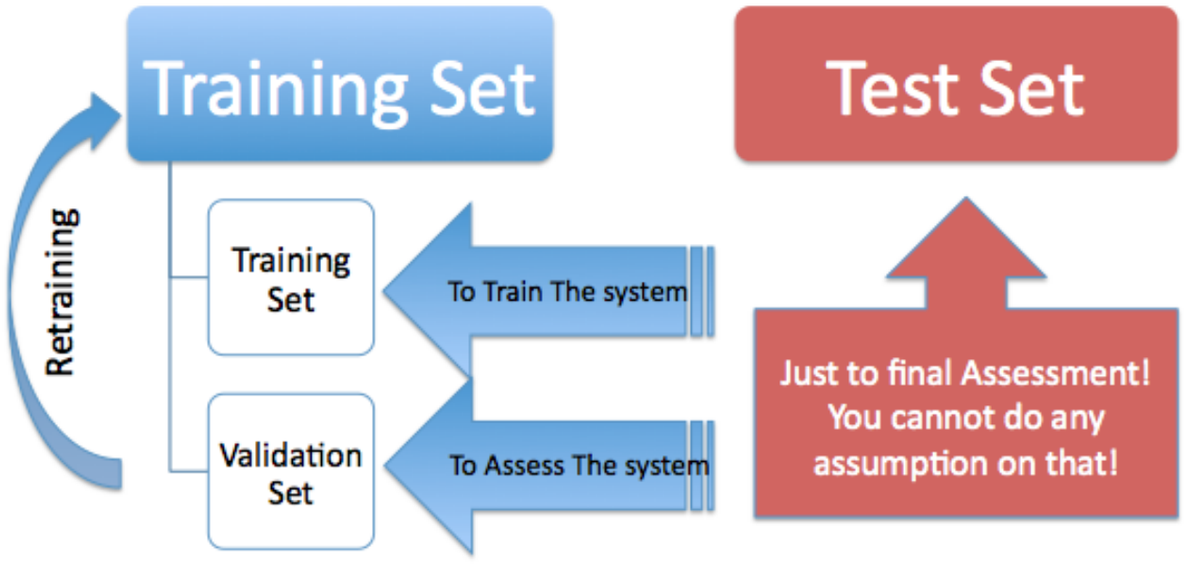
Learning Process– General Schema



Learning Process– General Schema

Steps:

1. Training Set:
   - Feature extraction
   - Data Modelization
2. Validation Set
   - Optimization of the model
3. Iterate 1 and 2
4. Test Set:
   - Final assessment!
   - No assumption using these data

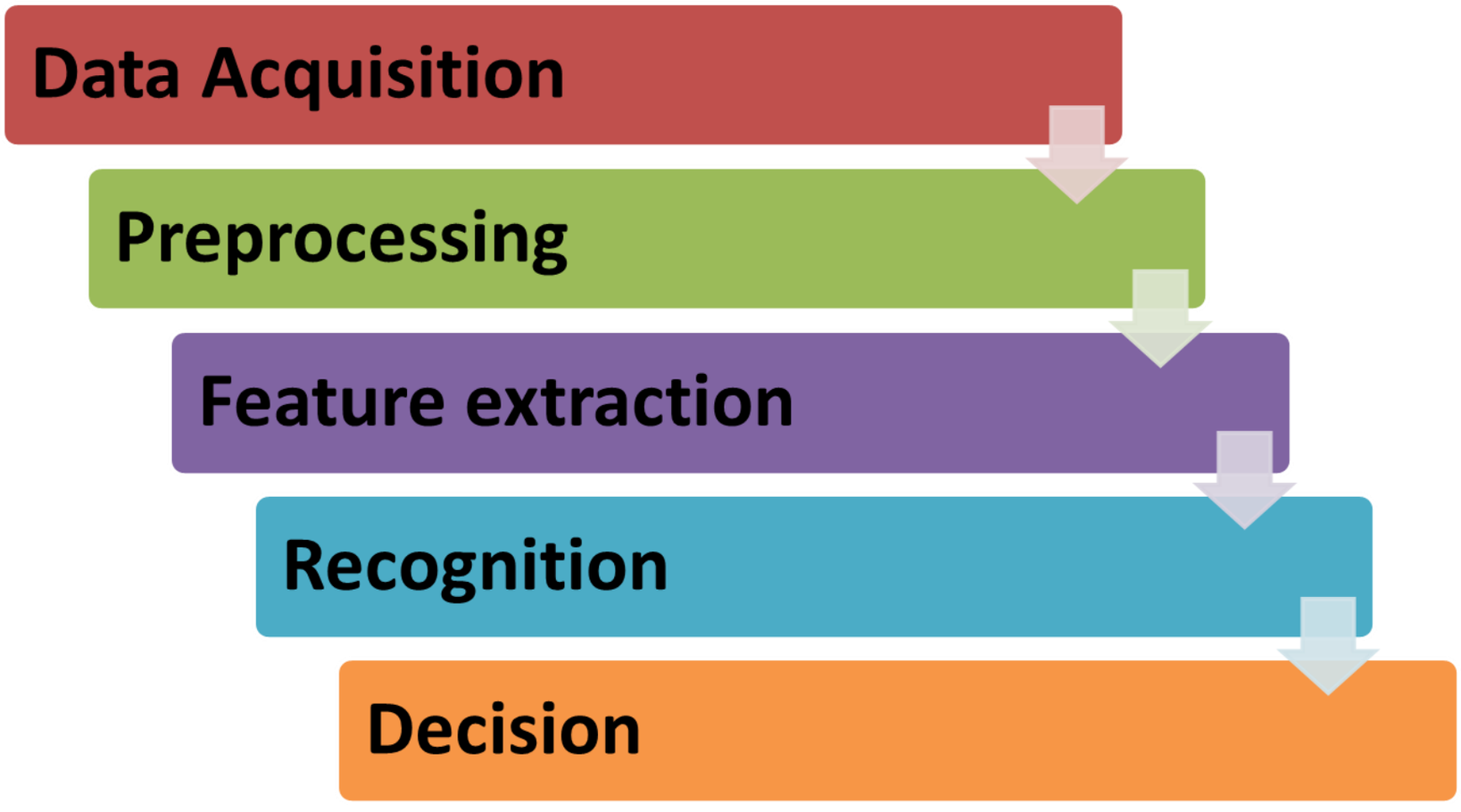


http://textanddatamining.blogspot.ch/2011/09/how-classifier-accuracy-is-conditioned.html

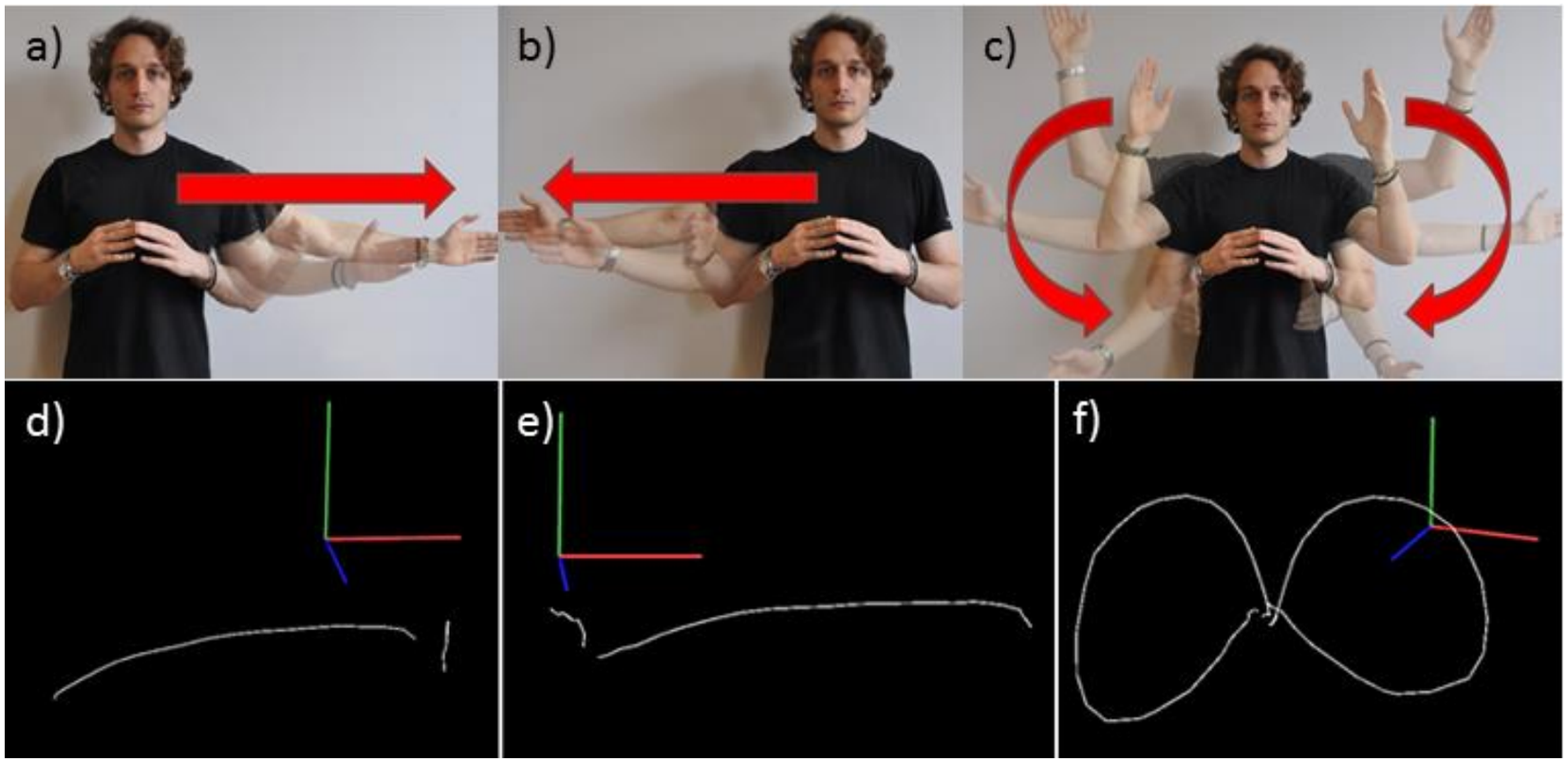
Learning Process– General Schema

- Dataset
  - **Training set:**
    - A set of examples used for learning, where the target value is known.
  - **Validation set:**
    - A set of examples used to tune the architecture of a classifier and estimate the error.
  - **Test set:**
    - Used only to assess the performances of a classifier.
    - It is never used during the training process so that the error on the test set provides an unbiased estimate of the generalization error.

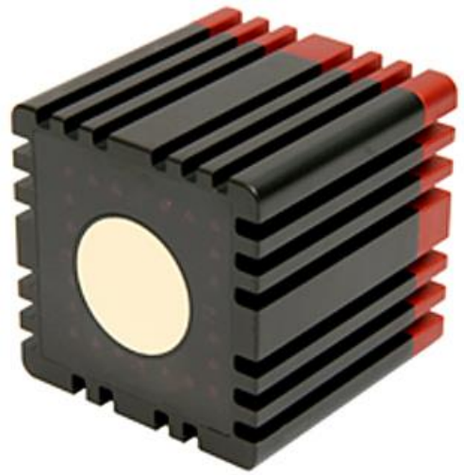
Information processing – general schema

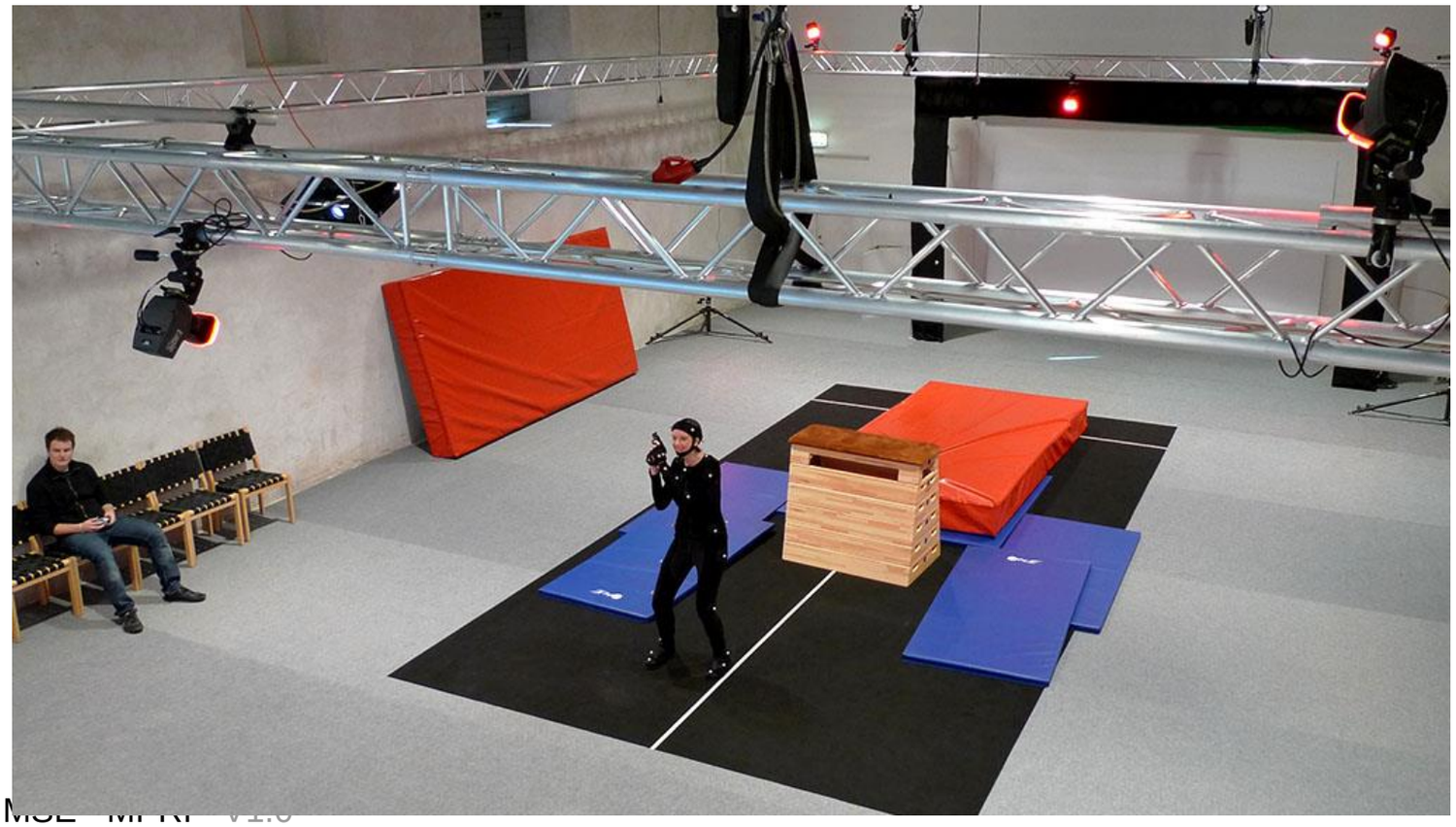


Example



Data Acquisition





- Driver
- Networking
- Synchronization
- …

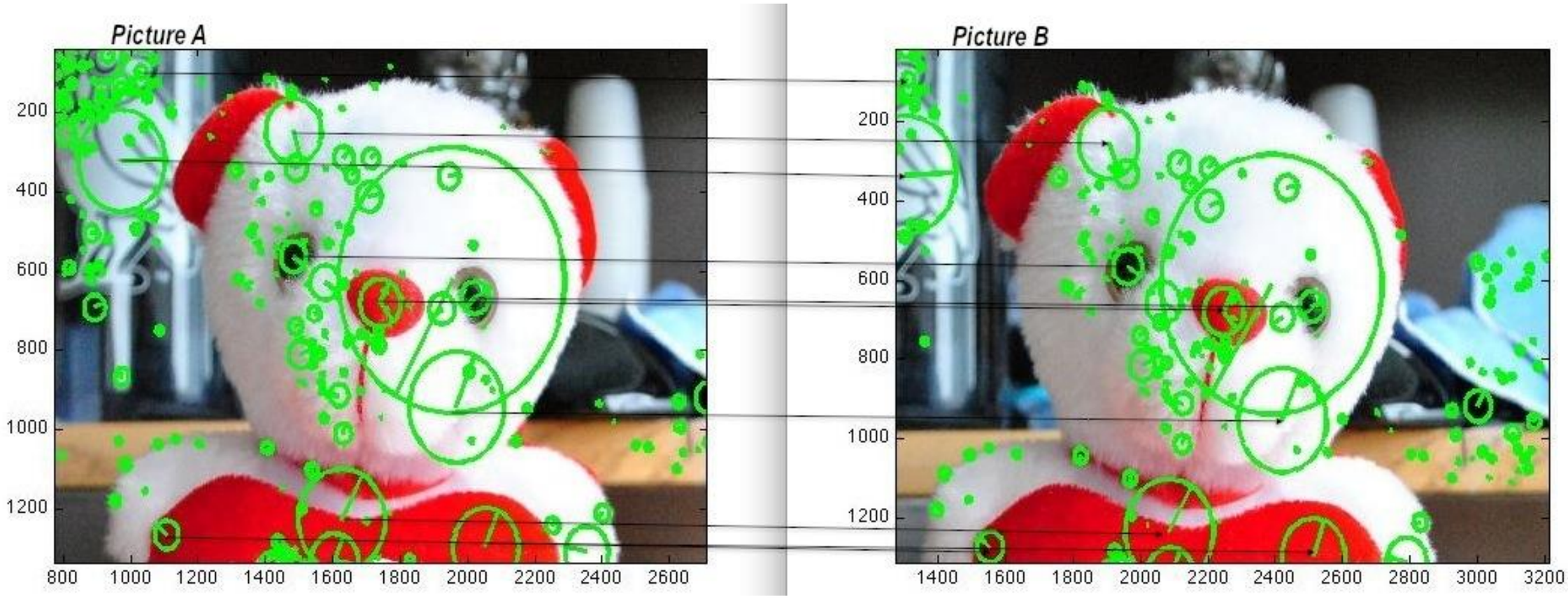
Preprocessing

- Data normalization:
  - Frame resizing
- Data segmentation:
  - Color segmentation
    - ✓ Hand detection
    - ✓ Color marker detection
- Motion segmentation:
  - Background subtraction
    - Works good on known background (static background)
    - Cannot detect stationary hands or determine which moving object is the hand

Feature Extraction

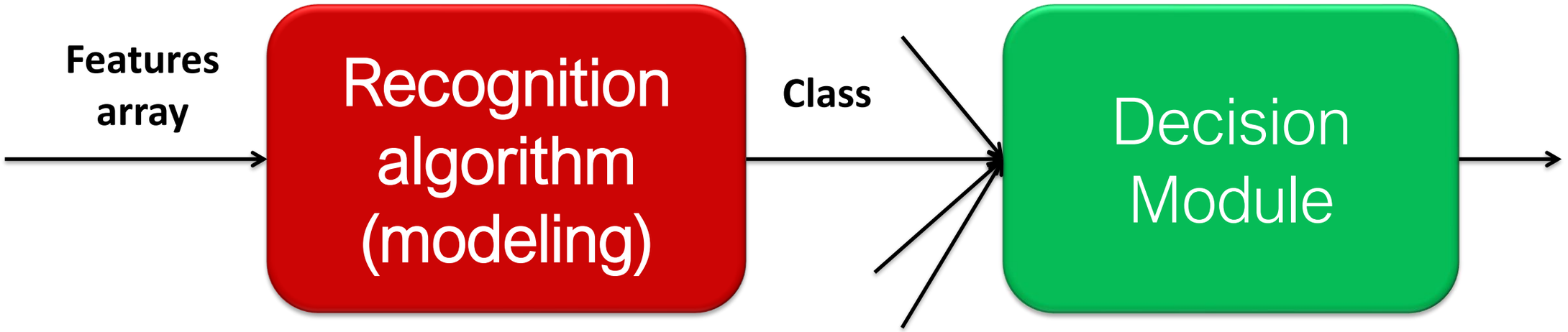
- Essential step in pattern recognition and machine learning problems.
- It is often decomposed into **feature construction** and **feature selection**
  - Feature Construction:
    - Goal: balance expressiveness of features with the size of the corresponding feature space
  - Feature Selection:
    - Goal: Select relevant and informative features

Feature Extraction



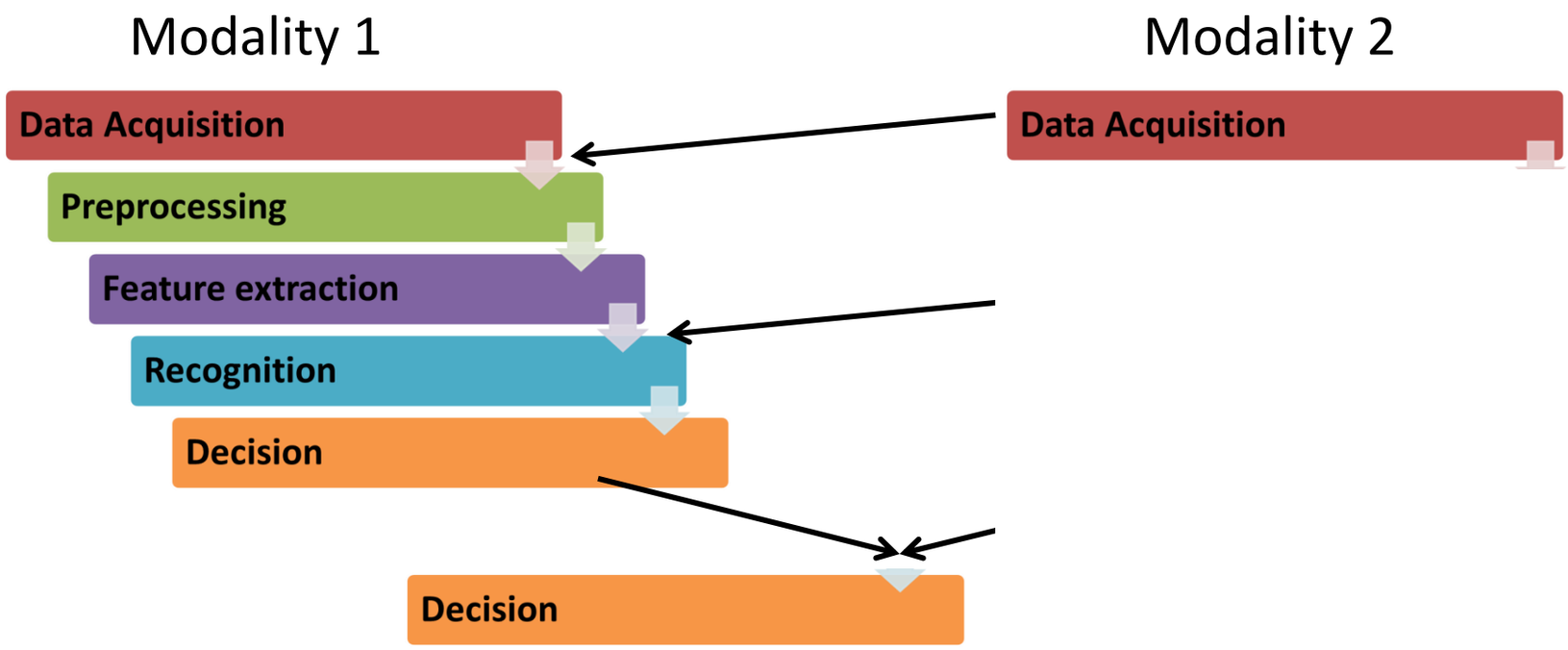
- SIFT descriptor: 128 elements

Recognition & Decision



- Decision Trees, Hidden Markov Model (HMM), Support vector machine, (Deep) Neural Networks
- Linear/Polynomial regression, Logistic regression, Conditional Random Fields (CRF), K-nearest neighbor (KNN), Dynamic time warping (DTW), boosting algorithms (AdaBoost, CatBoost, etc.), etc.
- Hybrid algorithms

Multimodal system – Information fusion



Information processing – general schema

Data Acquisition
- Sensors drivers
- Synchronization

Preprocessing
- Filtering, cropping, dropping, denoising, …
- Segmentation

Feature extraction
- Feature construction (selection)
- Feature reduction

Recognition
- Supervised approaches
- Unsupervised approaches

Decision
- Fusion

Conclusion

- Multimodality (definition)
- Machine Learning
  - Supervised Vs Unsupervised Learning
  - Classification Vs Regression
- Learning Process
  - Training set Vs Validation Set Vs Test Set
- Information processing
    - Data acquisition, pre-processing, feature extraction, recognition, decision