Introduction to Reinforcement Learning

2025. 1st semester

Prof. Taeyoon Kim

- *Affiliation: Dept. of Mobile System Engineering, Dankook Univ.*
- *Education*
  - *Received B.S., M.S., and Ph.D. in Dept. of Electrical & Electronic Engineering, Korea Univ.*
- *Career*
  - *Researcher in R&D center of Hyundai motors company*
- *Research Interests*
  - *Mobile Network and Communication tech.*
  - *Connected Swarm Intelligence*
- *Contact*
  - *301 in International Studies building (국제관)*
  - *Email: 2000kty@dankook.ac.kr*

Lecture Plan

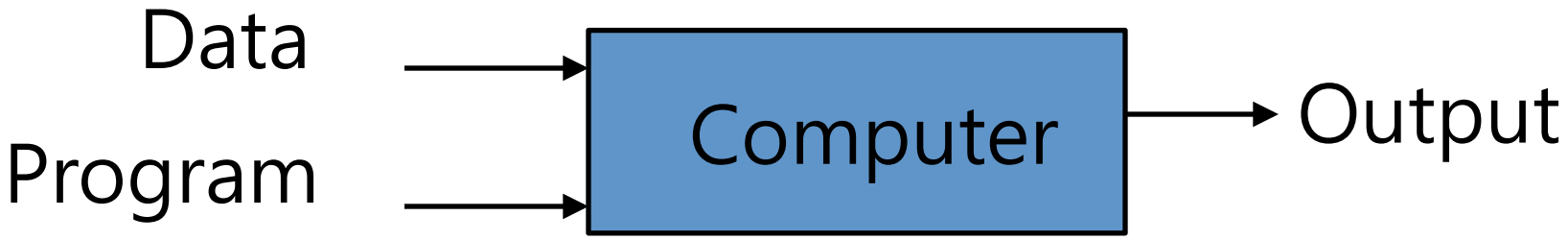
- *Lecture material*
  - *Lecture slides will be provided*
- *Performance evaluation*
  - *Attendance : 10%*
  - *Assignment : 10%*
  - *Exam : 80% (Mid-term: 40, Final: 40)*
- *Lecture Method*
  - *offline*

What is Machine Learning?

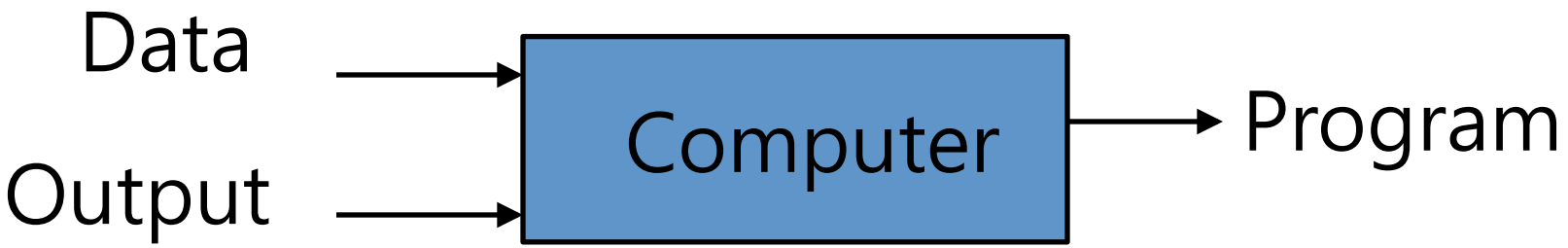
- "*Learning is any process by which a system improves performance from experience.*" -- *Herbert Simon*

- *Definition by Tom Mitchell (1998):*
  - *Machine Learning is the study of algorithms that*
    - *improve their performance P*
    - *at some task T*
    - *with experience E.*
  - *A well-defined learning task is given by <P, T, E>.*

Traditional Programming



Machine Learning

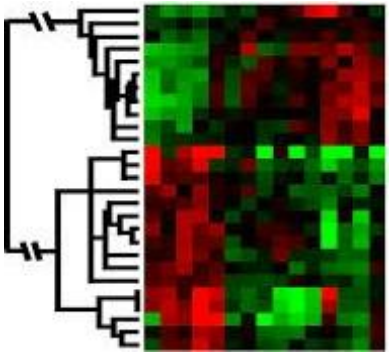


Slide credit: Pedro Domingos

When Do We Use Machine Learning?

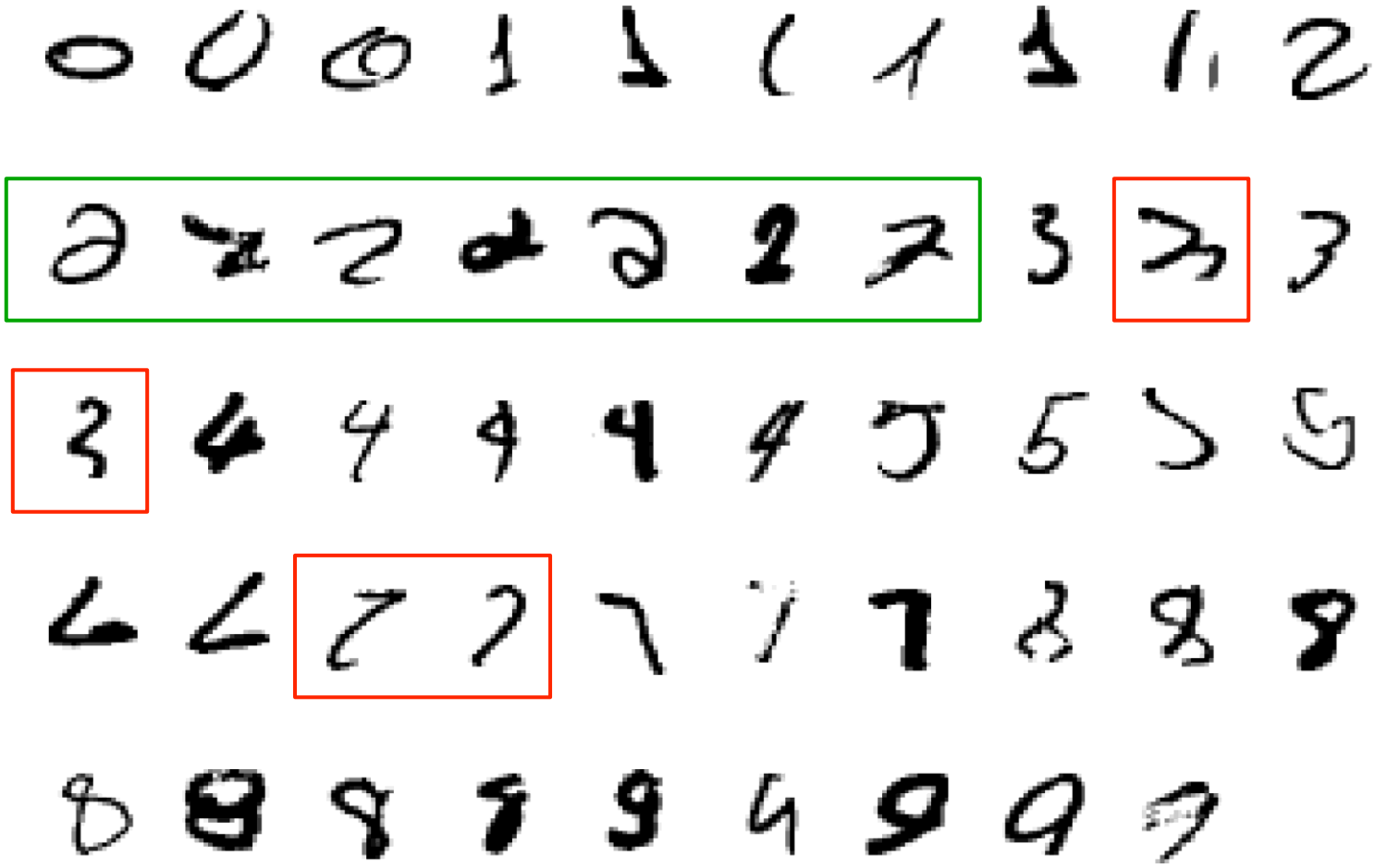
- ML is used when:
  - Human expertise does not exist (navigating on Mars)
  - Humans can't explain their expertise (speech recognition)
  - Models must be customized (personalized medicine)
  - Models are based on huge amounts of data (genomics)



- Learning isn't always useful:
  - There is no need to "learn" to calculate payroll

A classic example of a task that requires machine learning: It is very hard to say what makes a 2



Slide credit: Geoffrey Hinton

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 paderns of sensor readings in a nuclear power plant
- **Prediction:**
  - Future stock prices or currency exchange rates

Sample Applications

- *Web search*
- *Computational biology*
- *Finance*
- *E-commerce*
- *Space exploration*
- *Robotics*
- *Information extraction*
- *Social networks*
- *Debugging software*
- *[Your favorite area]*

Samuel's Checkers-Player

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

Dfining the Learning Task

Improve on task T,
with respect to performance metric P,
based on experience E

- *T: Playing checkers*
- *P: Percentage of games won against an arbitrary opponent*
- *E: Playing practice games against itself*

- *T: Recognizing hand-written words*
- *P: Percentage of words correctly classified*
- *E: Database of human-labeled images of handwritten words*

- *T: Driving on four-lane highways using vision sensors*
- *P: Average distance traveled before a human-judged error*
- *E: A sequence of images and steering commands recorded while observing a human driver*

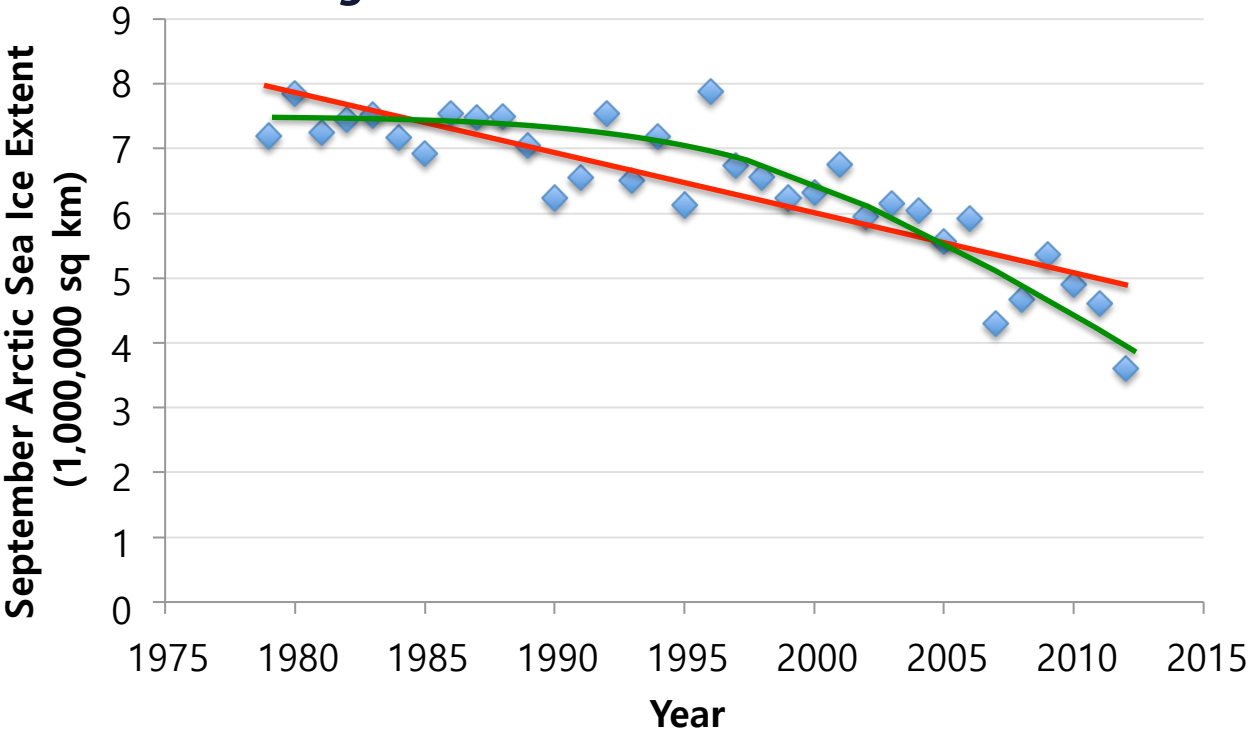
- *T: Categorize email messages as spam or legitimate*
- *P: Percentage of email messages correctly classified*
- *E: Database of emails, some with human-given labels*

Types of Learning

- *Supervised (inductive) learning*
  - *Given: training data + desired outputs (labels)*
- *Unsupervised learning*
  - *Given: training data (without desired outputs)*
- *Semi-supervised learning*
  - *Given: training data + a few desired outputs*
- *Reinforcement learning*
  - *Rewards from sequence of actions*

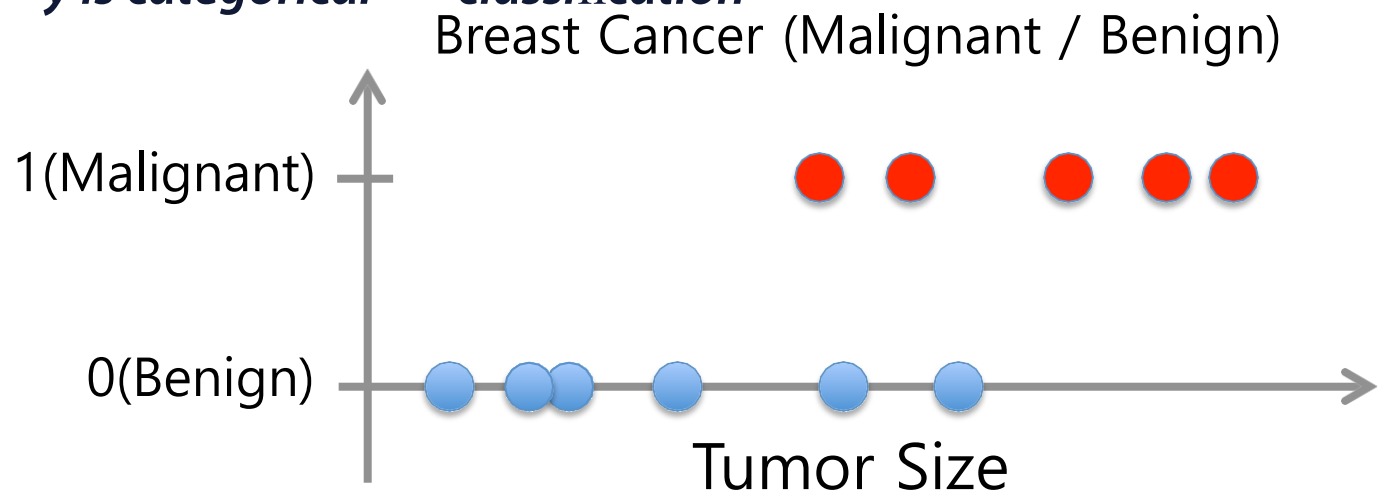
Supervised Learning: Regression

- *Given (x1, y1), (x2, y2), ..., (xn, yn)*
- *Learn a function f(x) to predict y given x*
  - *y is real-valued == regression*



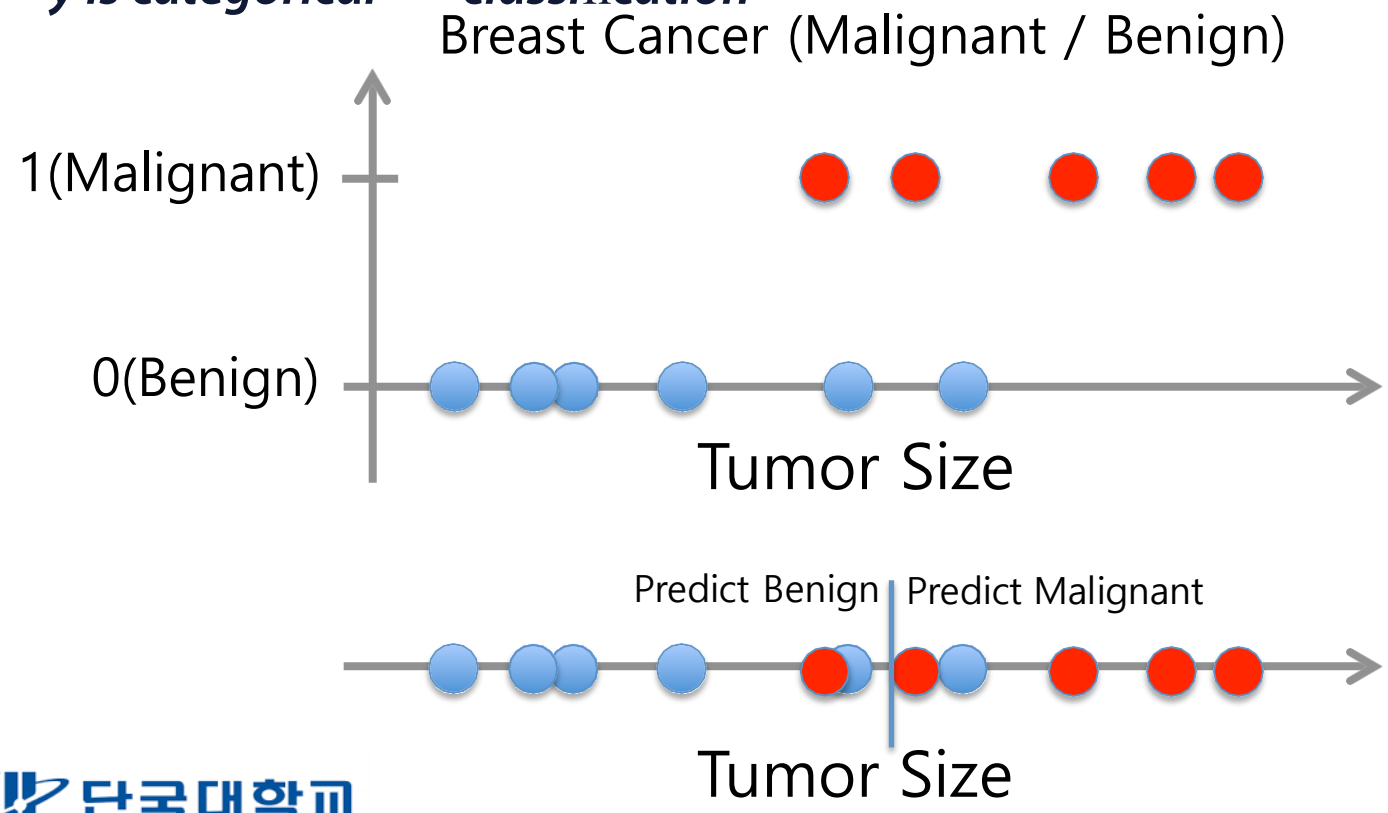
Supervised Learning: Classification

- *Given (x1, y1), (x2, y2), ..., (xn, yn)*
- *Learn a function f(x) to predict y given x*
  - *y is categorical == classification*



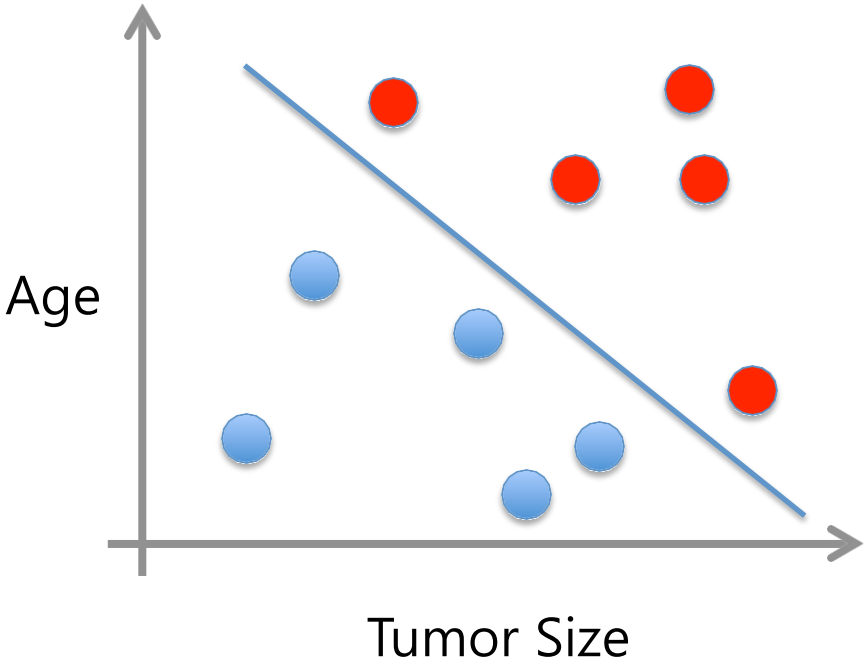
Supervised Learning: Classification

- *Given (x1, y1), (x2, y2), ..., (xn, yn)*
- *Learn a function f(x) to predict y given x*
  - *y is categorical == classification*



Supervised Learning

- *x can be multi-dimensional*
  - *Each dimension corresponds to an attribute*

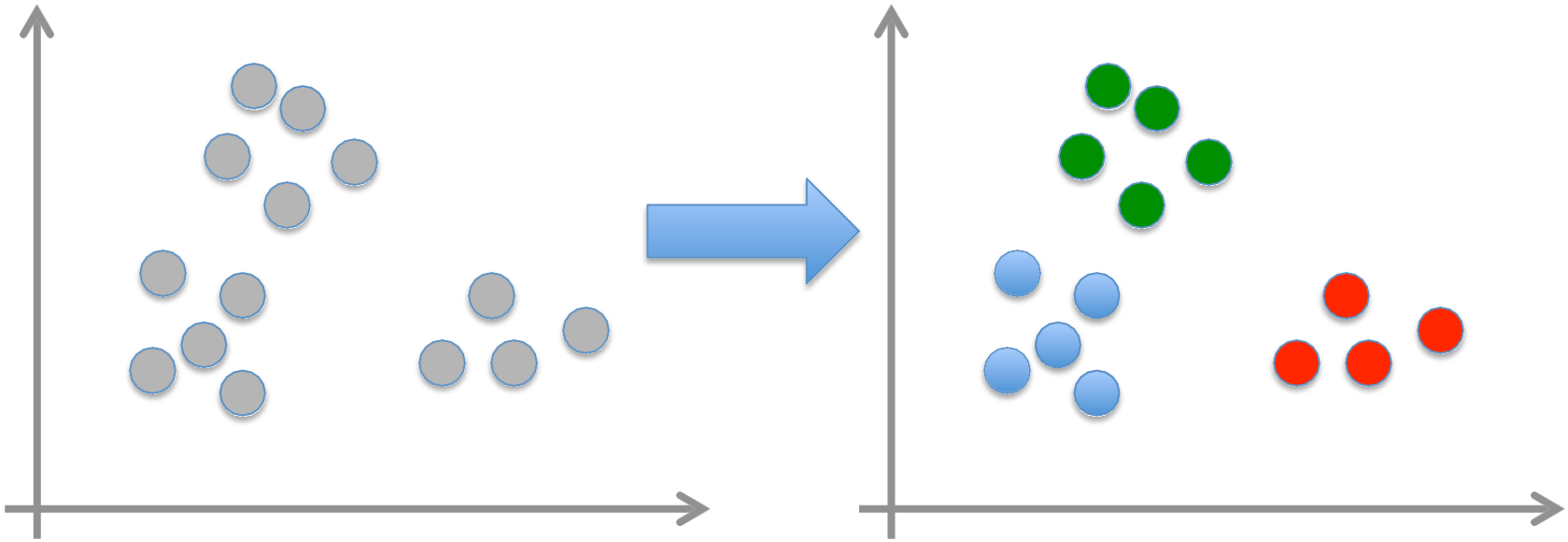


- Clump Thickness
- Uniformity of Cell Size
- Uniformity of Cell Shape

…

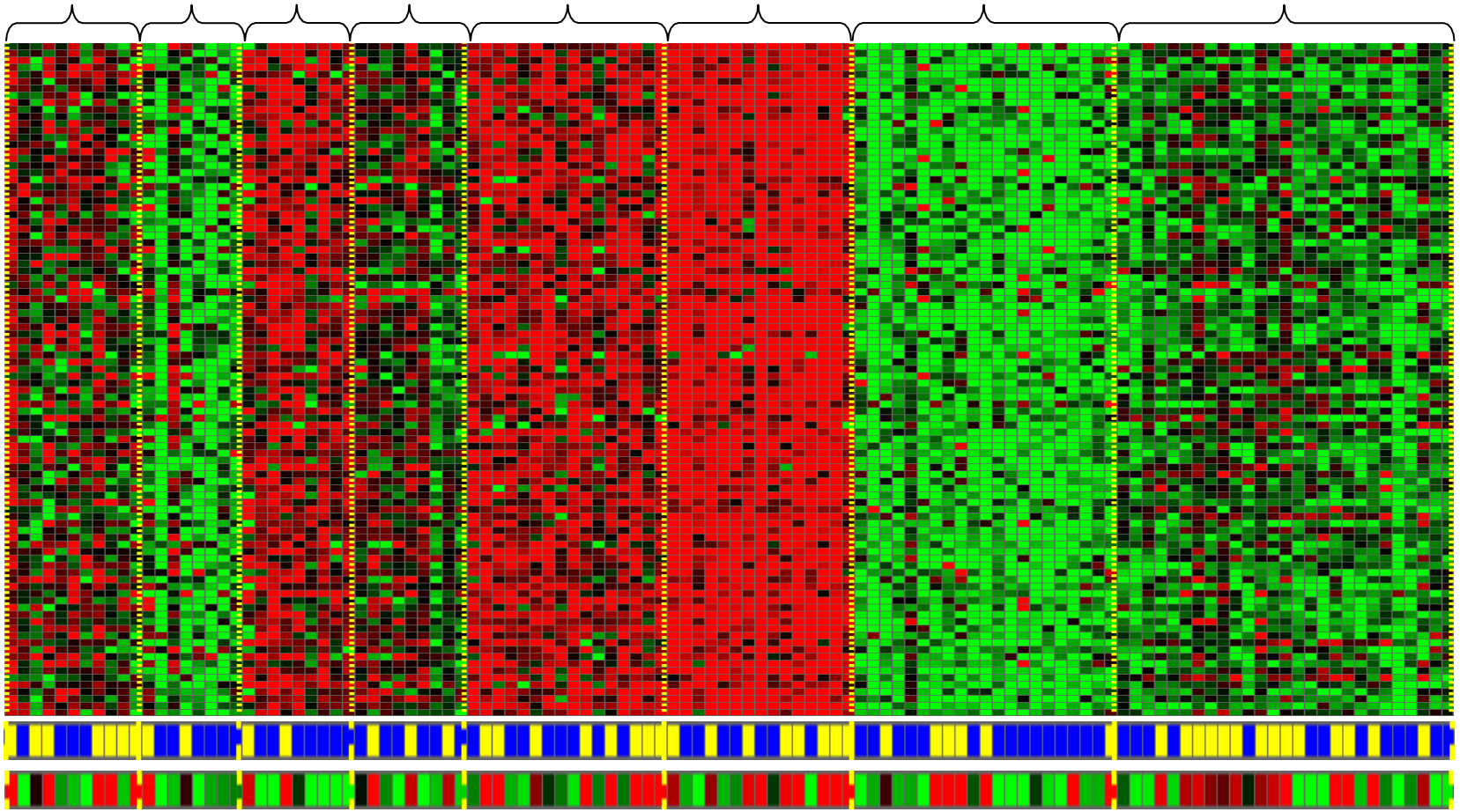
Unsupervised Learning

- *Given x1, x2, …, xn (without labels)*
- *Output hidden structure behind the x's*
  - *E.g., clustering*



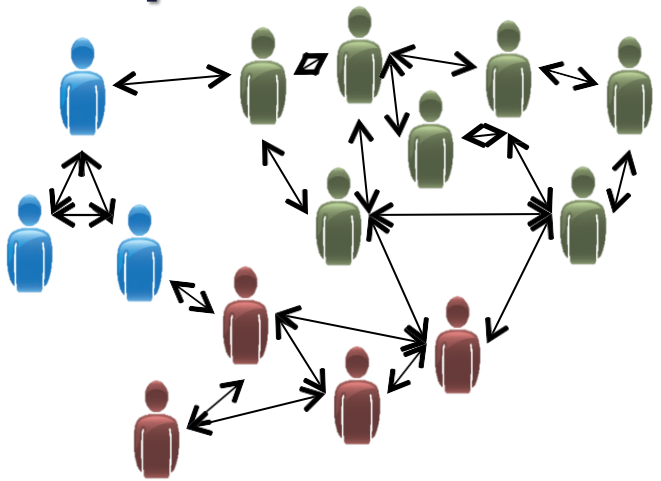
Unsupervised Learning

Genomics application: group individuals by genetic similarity



Unsupervised Learning

Organize computing clusters



Social network analysis

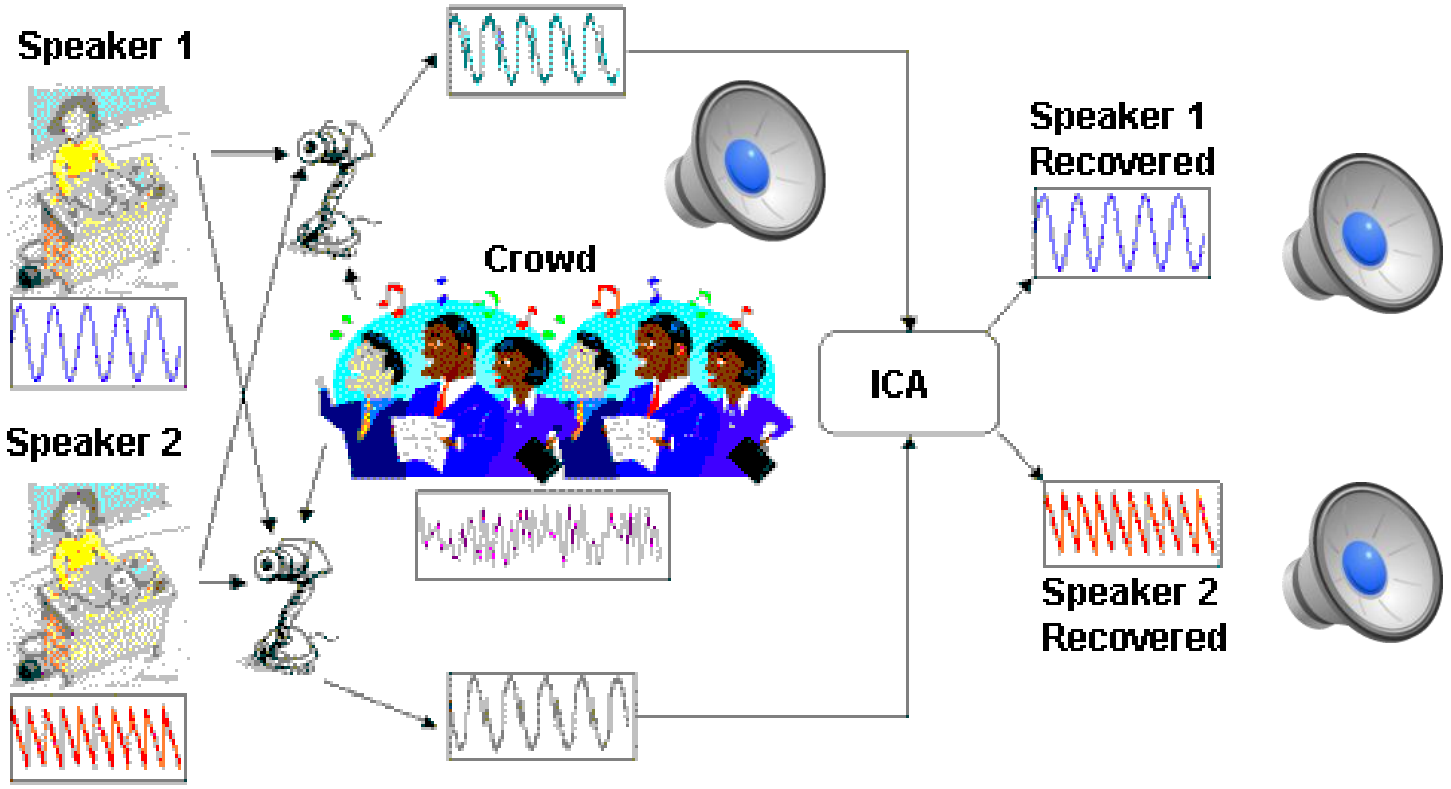
Market segmentation

Image credit: NASA/JPL-Caltech/E. Churchwell (Univ. of Wisconsin-Madison)

Astronomical data analysis

Unsupervised Learning

- Independent component analysis
  - separate a combined signal into its original sources

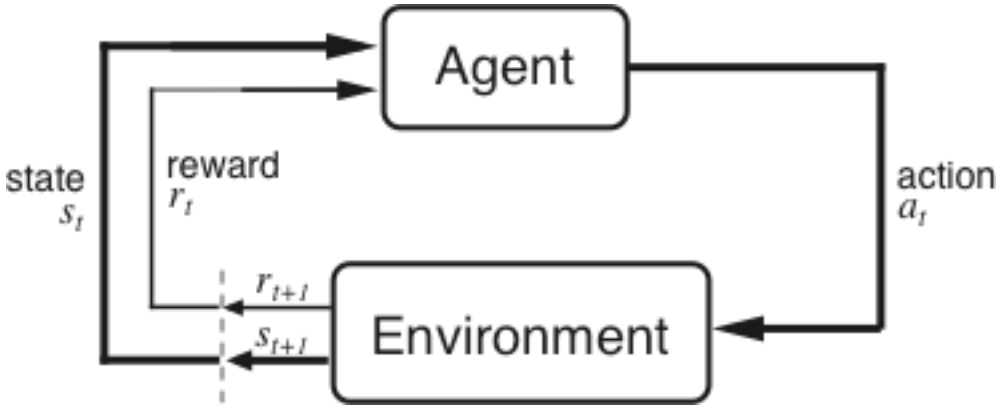


Reinforcement Learning

- *Given a sequence of states and actions with (delayed) rewards, output a policy*
  - *Policy is a mapping from states → actions that tells you what to do in a given state*

- *Examples:*
  - *Credit assignment problem*
  - *Game playing*
  - *Robot in a maze*
  - *Balance a pole on your hand*

The Agent-Environment Interface



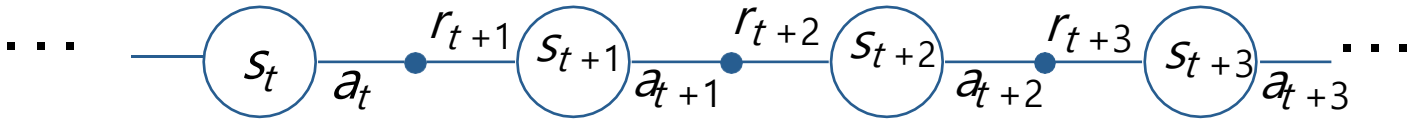
Agent and environment interact at discrete time steps : $t = 0, 1, 2, \mathrm{K}$

Agent observes state at step t : $s_t \in S$

produces action at step t : $a_t \in A(s_t)$

gets resulting reward : $r_{t+1} \in \Re$

and resulting next state : s_{t+1}



Reinforcement Learning

https://youtu.be/ldXxDNjS5jw