

Pattern Information Processing¹ (パターン情報処理)

Masashi Sugiyama
(Department of Computer Science)
杉山 将 (計算工学専攻)

Contact: W8E-505

sugi@cs.titech.ac.jp

<http://sugiyama-www.cs.titech.ac.jp/~sugi>

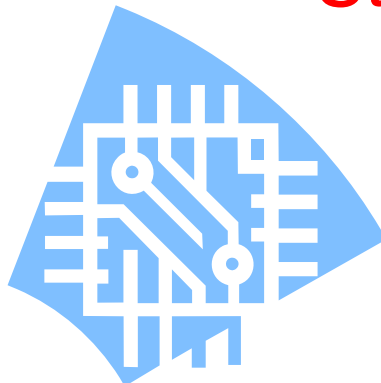
Contents of This Lecture (1)

- Syllabus (what I will provide in this course):

Inferring an underlying input-output dependency from input and output examples is called **supervised learning**.

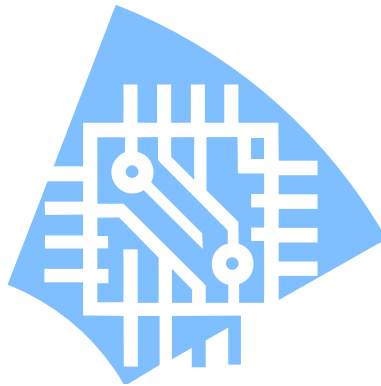
This course focuses on a statistical approach to supervised learning and introduces its basic concepts as well as state-of-the-art techniques.

Statistical machine learning

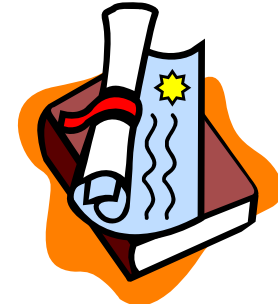


Contents of This Lecture (2)

- What you are expected to learn in this course:
 - How to use supervised learning methods
 - Ideas behind the methods
 - Novel research topics in supervised learning
 - Something useful in **your own research/life**



Grading



■ Small reports

- Almost every week
- Deadline: next class

■ Mini conference on supervised learning (final day)

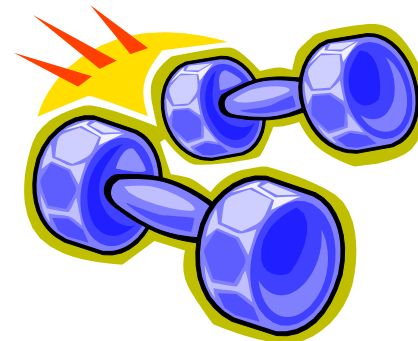
- Apply supervised learning techniques to your own data sets and analyze them!

■ Final reports on the above issue

Brief Overview of the Course (1)⁵

3 types of learning

- Supervised learning
- Unsupervised learning
- Reinforcement learning



Brief Overview of the Course (2)⁶

- Topics in supervised learning
 - Active learning
 - Model selection
 - Learning method

Textbook

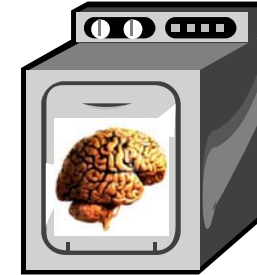
- Handouts are provided if necessary.
- Pointers to related articles will be provided.
- I suppose you know **elementary statistics** and **linear algebra**. If not, please study them by yourself!

3 Topics in Learning Research

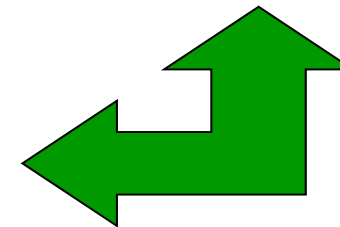
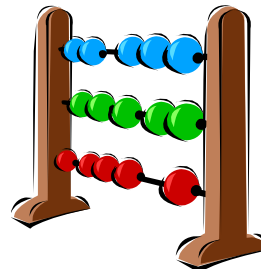
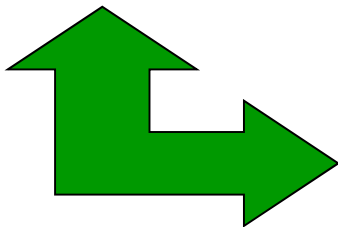
8



Understanding the brain
(Physiology, psychology,
neuroscience)



Developing learning machines
(Computer and electronic
engineering)



Mathematically clarifying mechanism of learning
(Computer and information science)

Three Types of Learning

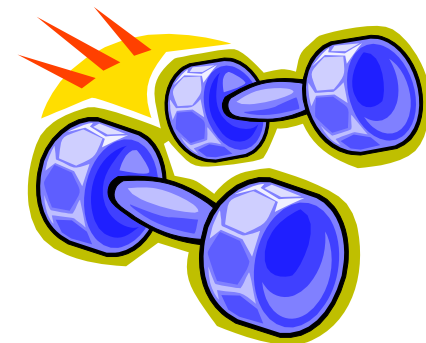
- Supervised learning
(This course!)



- Unsupervised learning

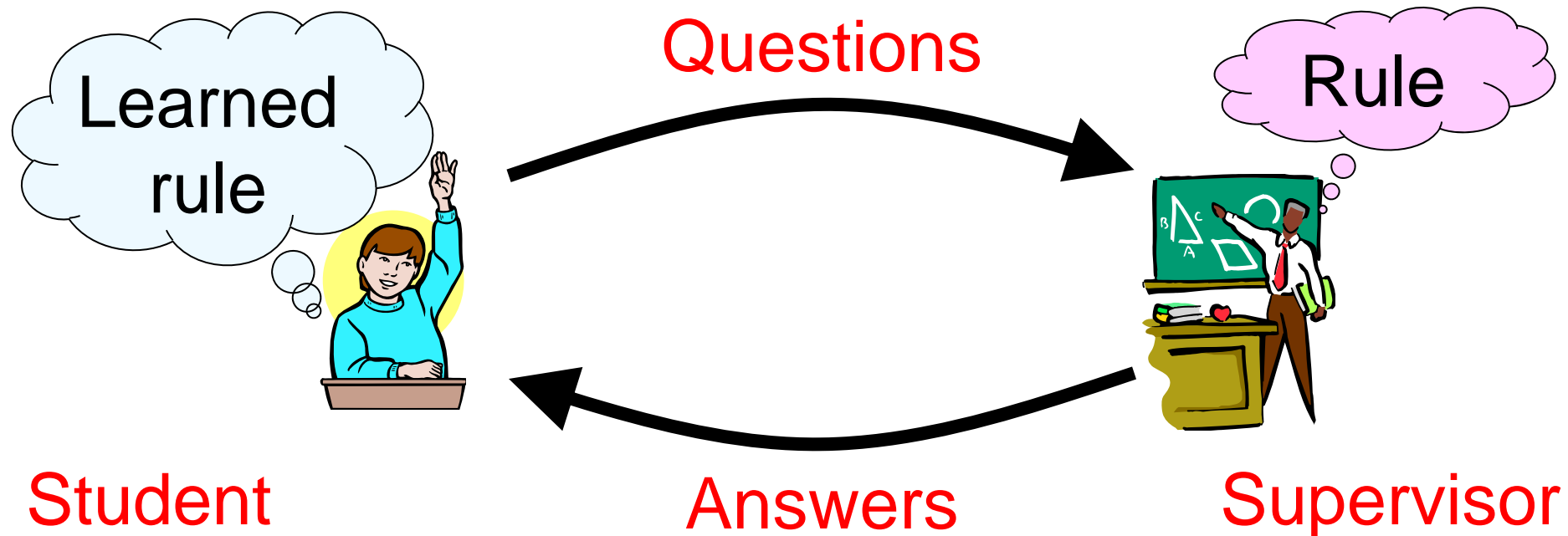


- Reinforcement learning



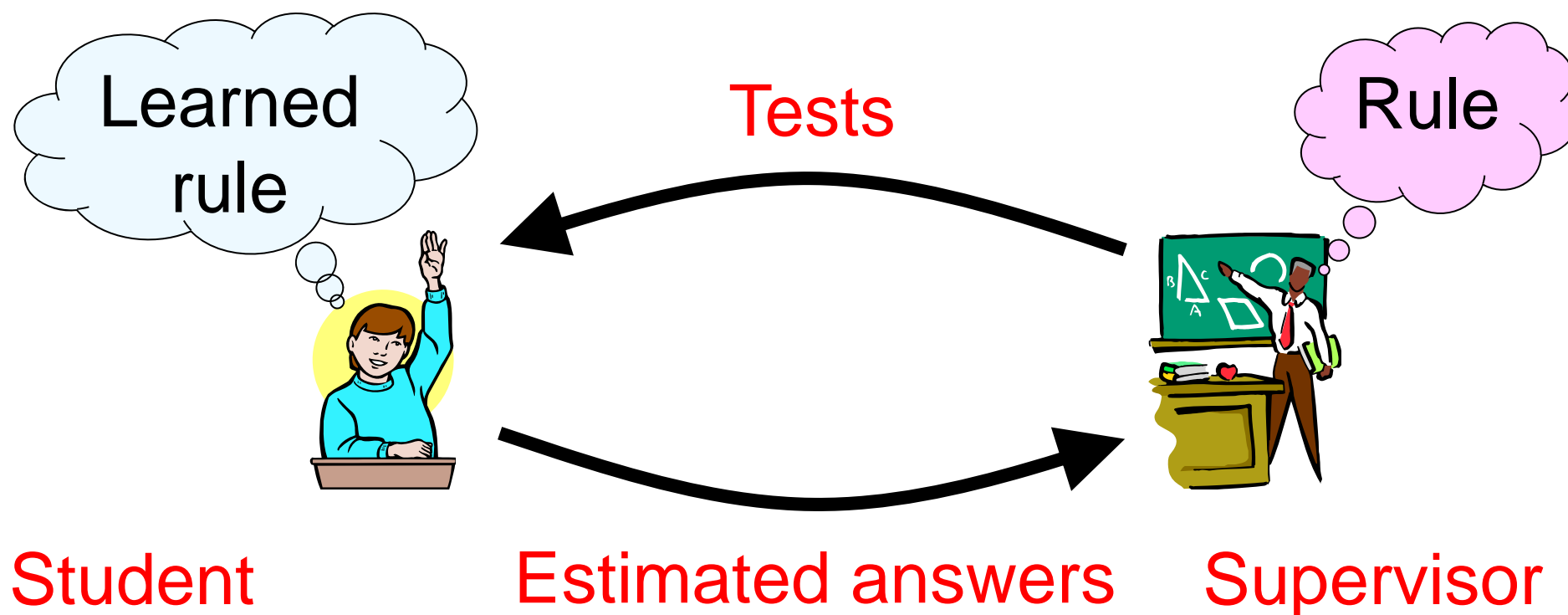
What Is Supervised Learning?¹⁰

- The goal of supervised learning is to estimate an **unknown input-output rule**.
- You are allowed to ask questions to a supervisor (“oracle”) who knows the rule.
- The supervisor answers your questions using the rule.



Generalization Capability

- **Training examples**: pairs of **questions** and **answers**.
- If the underlying rule can be successfully estimated, we can **answer to the questions that we have never taught**.
- Such an ability is called the **generalization capability**.



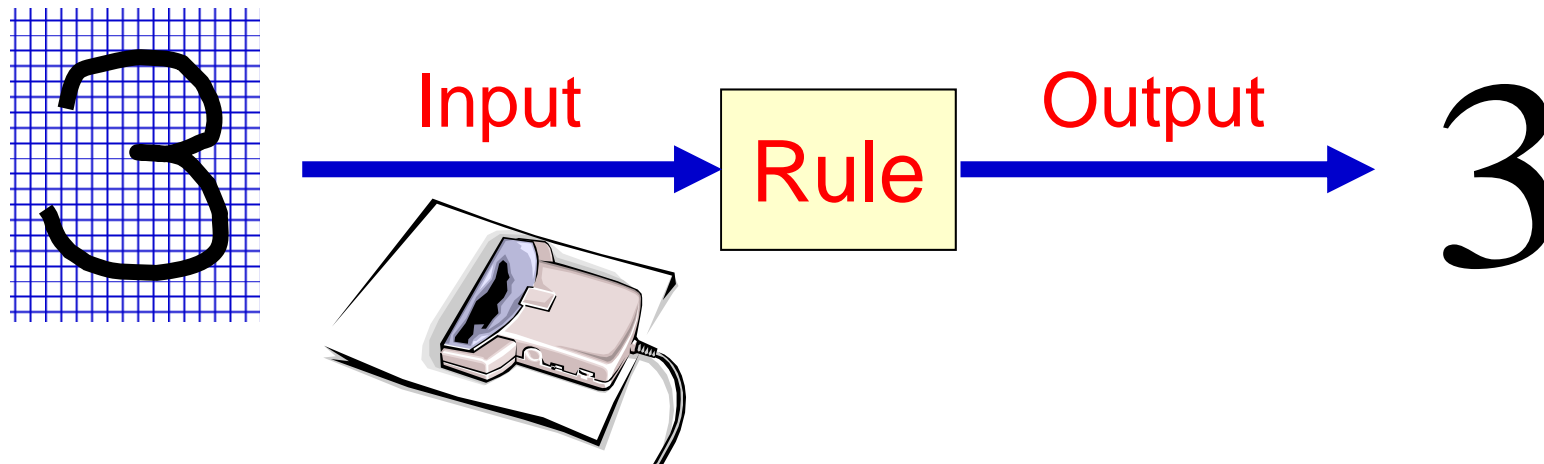
Example 1

- Hand-written number recognition

We want to recognize the scanned hand-written characters.

Example 1

- Training examples consist of { (hand-written number, its recognition result) }.
- If underlying input-output rule is successfully learned, unlearned hand-written numbers can be recognized.



Example 2:

Brain-Computer Interface (BCI)

- Control computers by brain signals:
 - **Input:** brain signals (EEG)
 - **Output:** Left or Right

Training

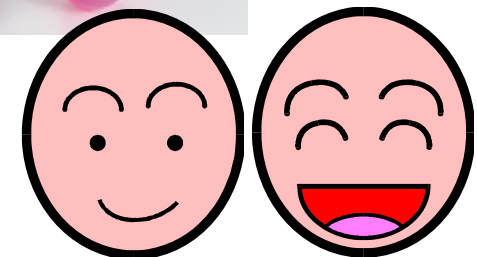
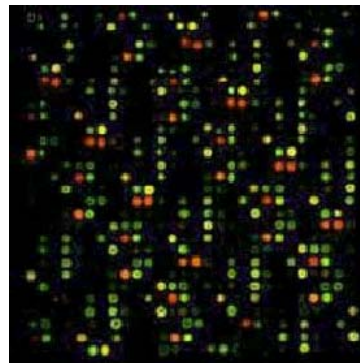
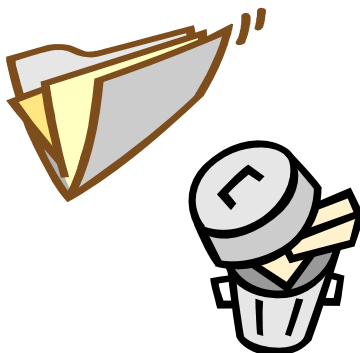
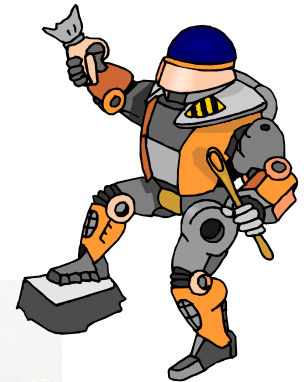
- Imagine left/right-hand movement following the letter on the screen

Testing: Playing Games

- “Brain-Pong”

Other Examples

- Other examples are...
 - Stock price estimation
 - Robot motor control
 - Computer vision
 - Spam filter
 - DNA classification

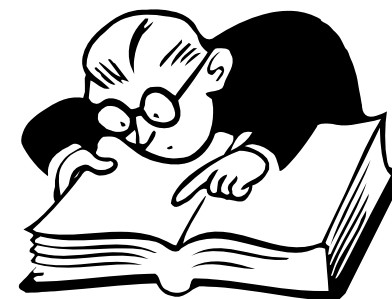


Three Types of Learning

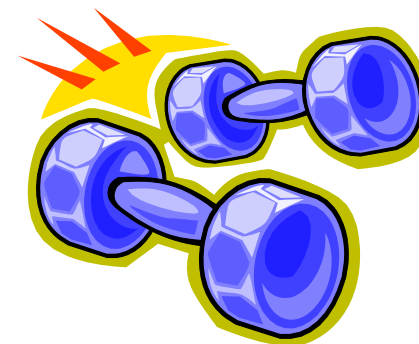
- Supervised learning
(This course!)



- Unsupervised learning
("Advanced data analysis",
odd year, spring)



- Reinforcement learning



What Is Unsupervised Learning?¹⁹

- You are given questions (input data) **without answers** (output data).
- The goal is to find “**interesting**” structure in the data.



Interestingness

- The goal of unsupervised learning depends on the definition of “interestingness”:
 - Dimensionality reduction
 - Clustering
 - Blind source separation
 - Outlier detection

Dimensionality Reduction

■ Dimensionality reduction (Embedding)

- We are given high-dimensional data.
- High-dimensional data is too complex to analyze: Even estimating the density is extremely difficult (“**curse of dimensionality**”)
- We want to have a low-dimensional expression of the data without losing **intrinsic** information.
- **Data visualization**: Reduced data is less than equal to 3-dimensional.

Example 1

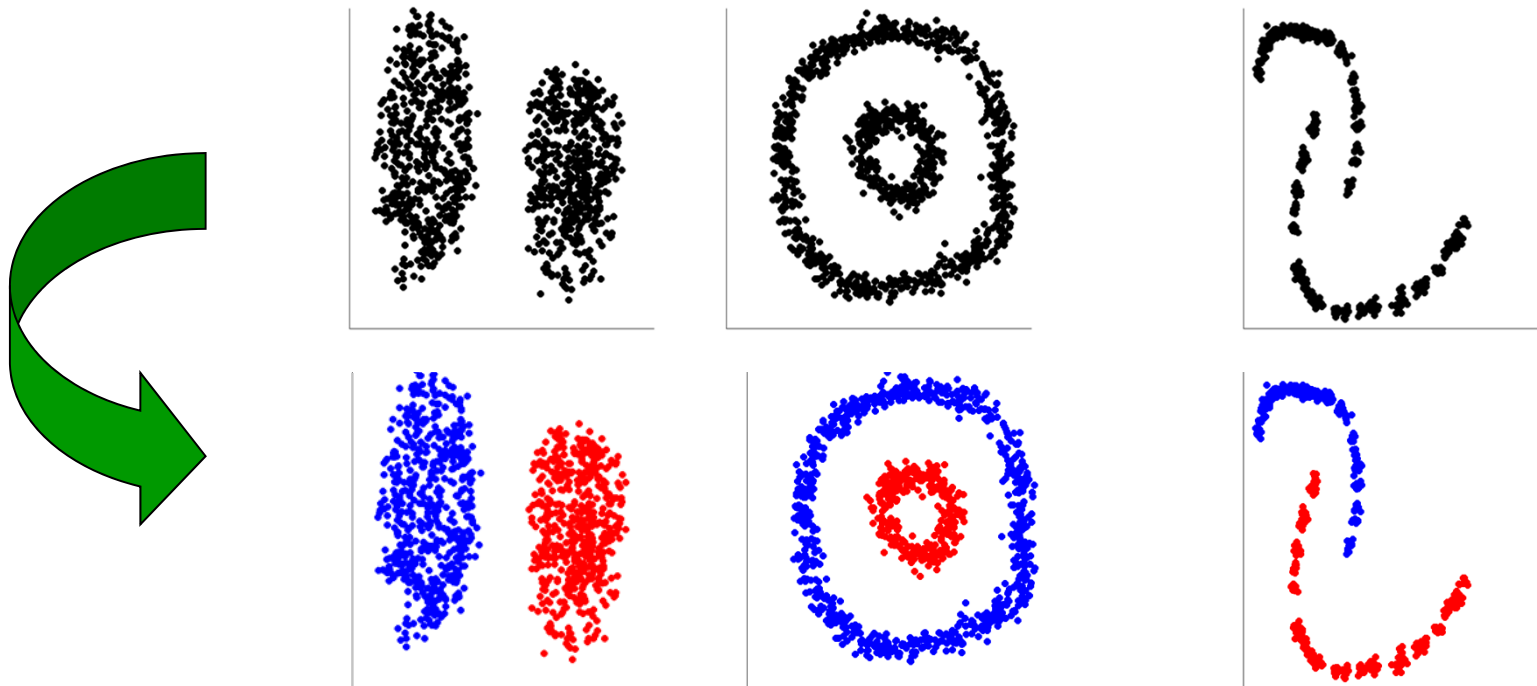
- “Swiss Roll”
- Data is 3-D but it essentially lies on a 2-D manifold.
- We want to “unfold” the roll.

Example 2

- Embedding face images into 2D space.
- Images of the same face from different angles and lighting directions (64x64=4096D)

Data Clustering

- We want to divide the data into disjoint groups:
 - Data in the **same group** are **similar**.
 - Data in **different groups** are **different**.
- “Unsupervised classification”

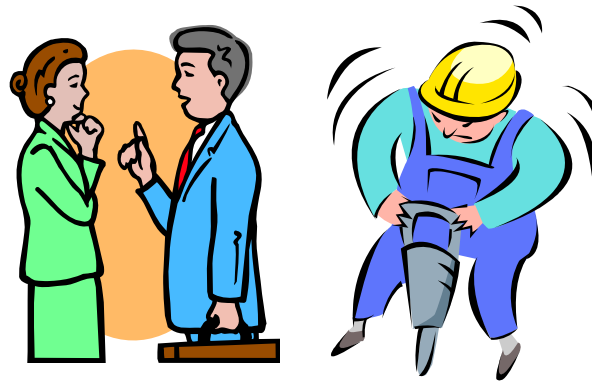


Example

- Image segmentation

Blind Source Separation

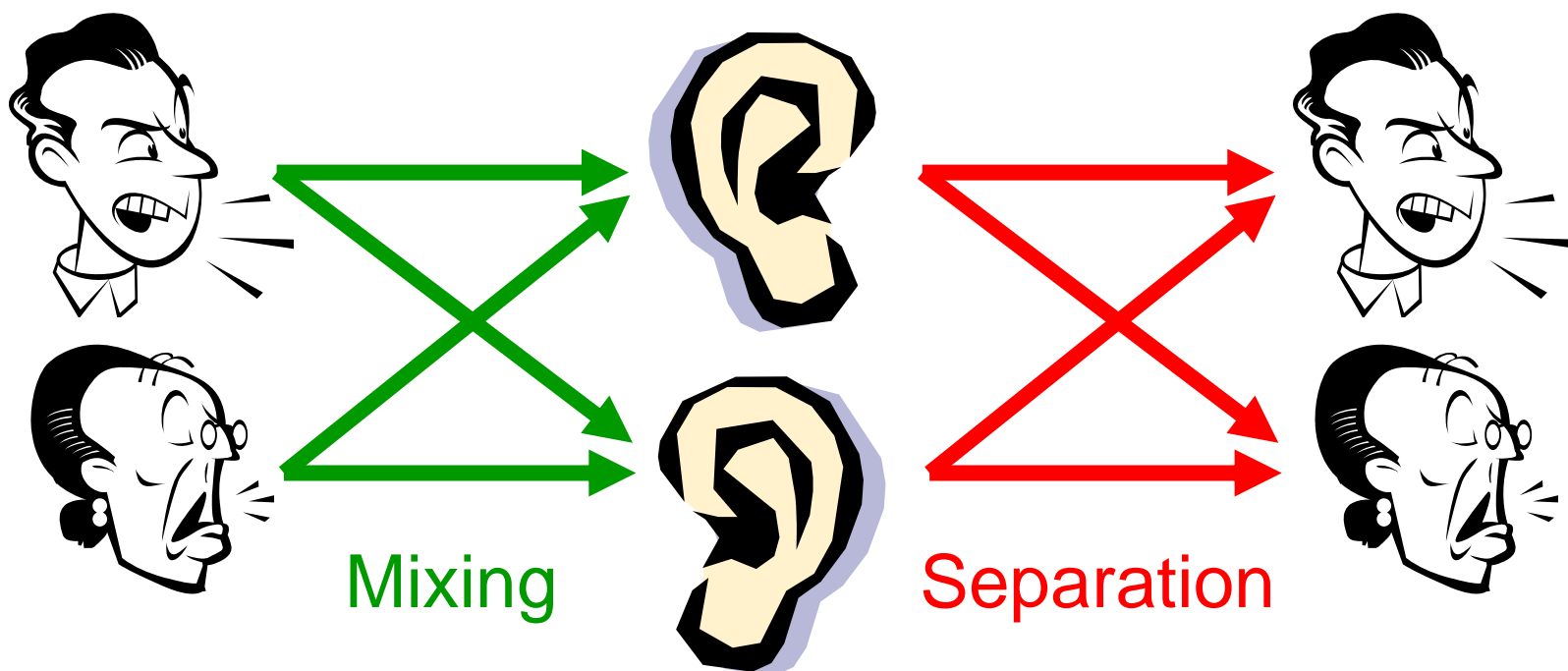
- We can extract what a person is speaking in a noisy environment.



- Syotoku-taishi can distinguish 10 conversations?

Blind Source Separation

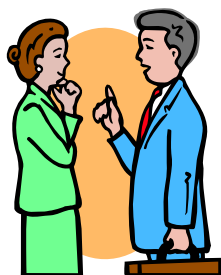
■ Cocktail-party problem



- We want to separate mixed signals into original ones.

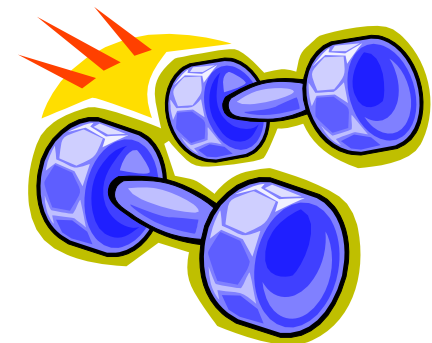
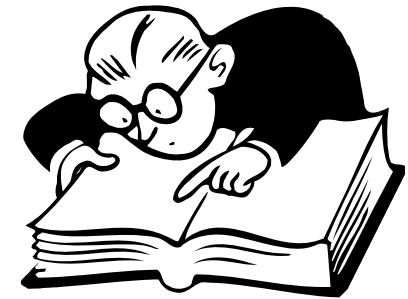
Example

	Mixed signal	Separated signal 1	Separated signal 2
Conversation + Conversation			
Conversation + Instrument			



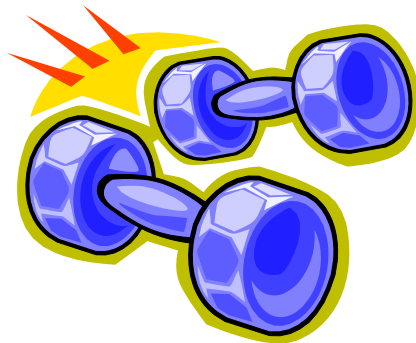
Three Types of Learning

- Supervised learning
(This course!)
- Unsupervised learning
("Advanced data analysis",
odd year, spring)
- Reinforcement learning
(to be covered...)



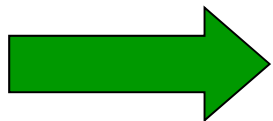
What Is Reinforcement Learning?³⁰

- The goal of reinforcement learning is **same as supervised learning**, i.e., to estimate an unknown underlying rule.
- However, different from supervised learning, we are **not allowed to ask questions** to the teacher.
- Instead, we can get **rewards (reinforcement signals)** for our estimated answer.
- Goal is to learn the rule with **maximum rewards**.

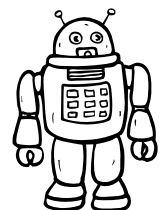


Example

- Learning stand-up motion
- The robot consists of 3 links connected by 2 joints.
- Robot can control its **joint angles** by itself.
- The goal is to learn the control rule for standing up.
- **Control rule**: mapping from inner states to control signal.
- Standing up = the head is at the highest position



Reward proportional to head's height



Example

- Before learning

Example

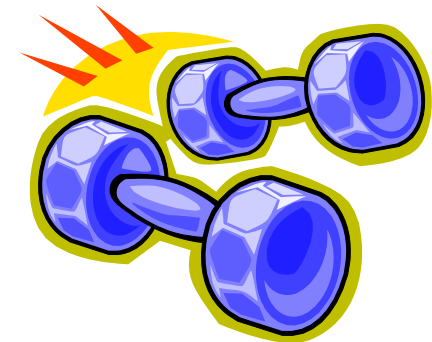
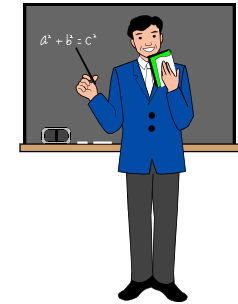
- After 750 trials

Example

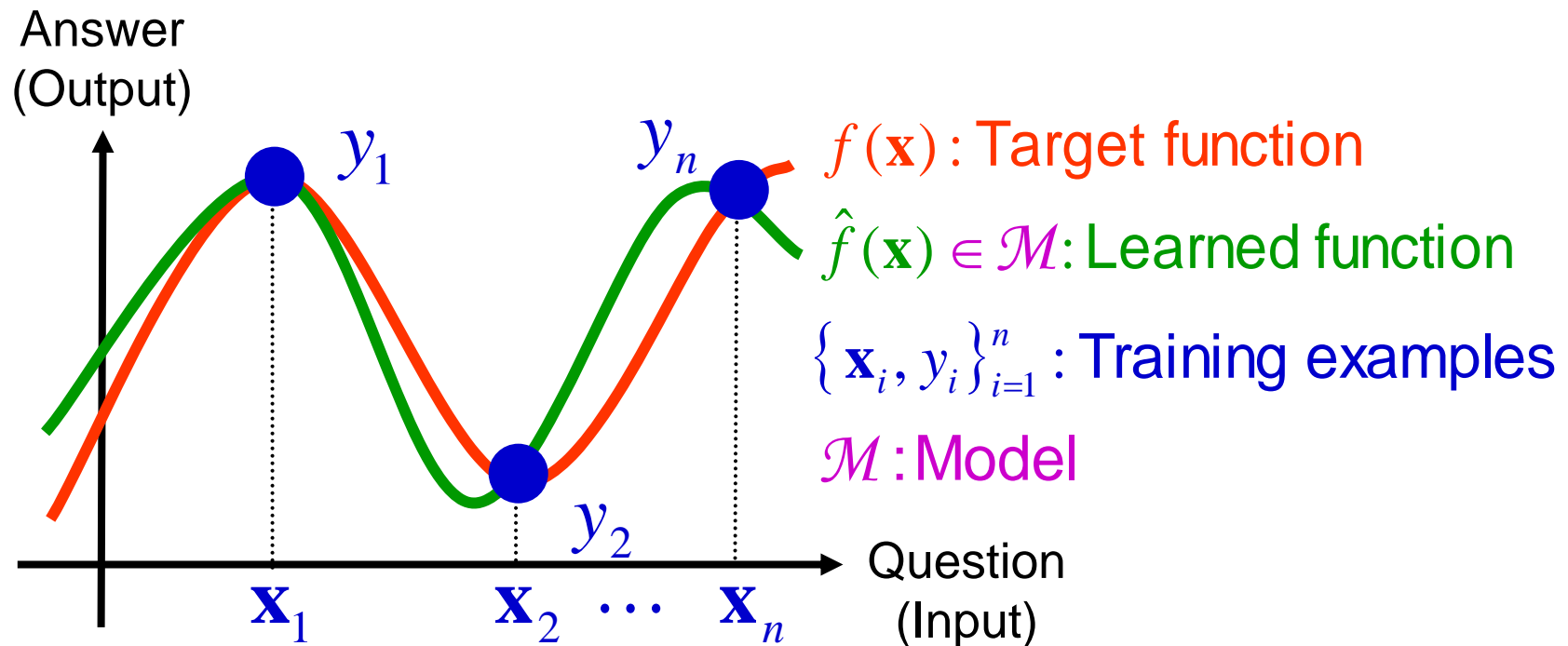
- After 920 trials

We Have Learned ...

- There are three issues in learning:
 - Supervised learning
 - Unsupervised learning
 - Reinforcement learning

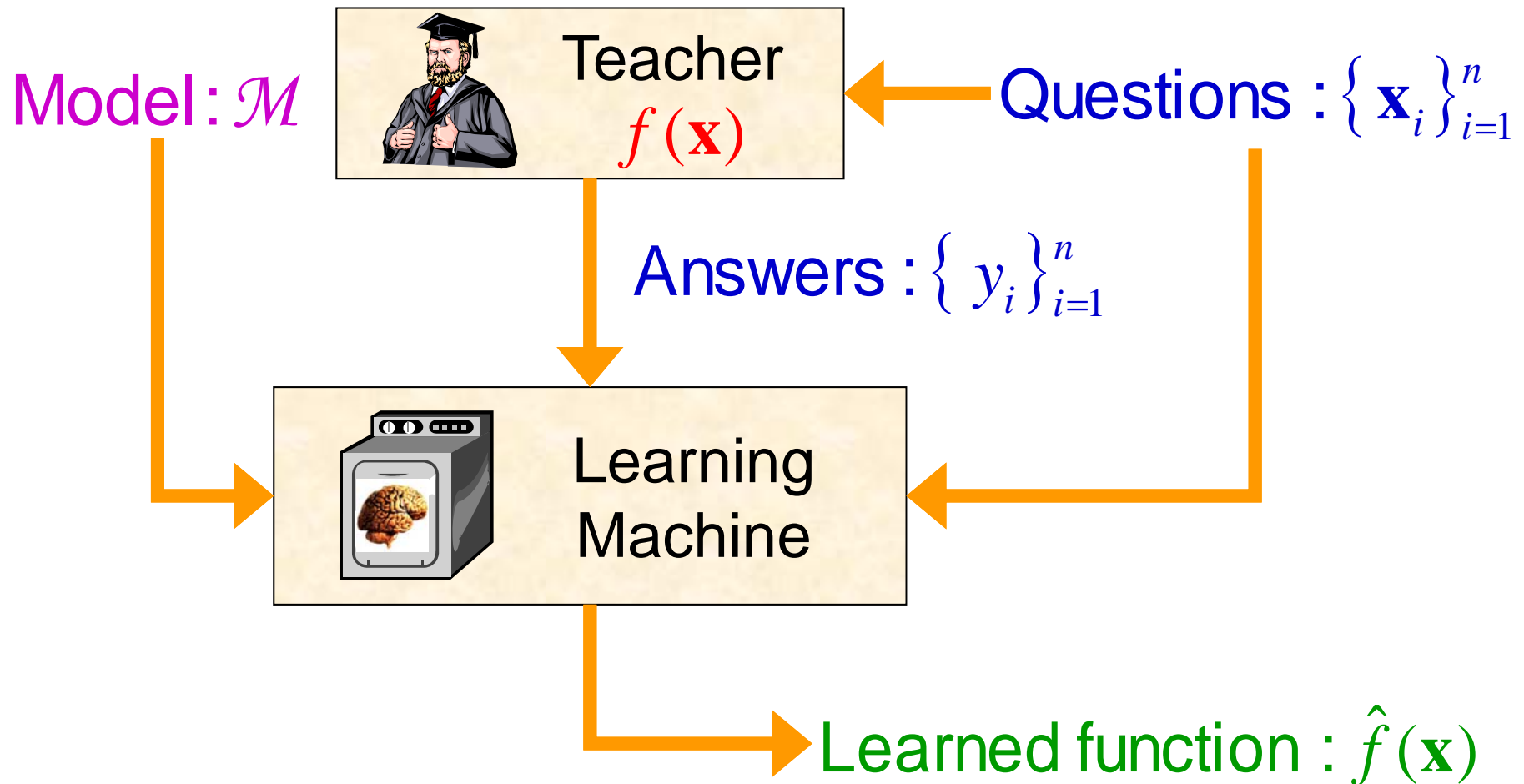


Supervised Learning As Function Approximation



Using training examples $\{\mathbf{x}_i, y_i\}_{i=1}^n$,
find a function $\hat{f}(\mathbf{x})$ from a model \mathcal{M}
that well approximates the target function $f(\mathbf{x})$.

Diagram of Supervised Learning³⁷



3 Important Topics in Supervised Learning

- **Active learning:**

What are the best questions to ask?

- **Model selection:**

What is the best model to use?

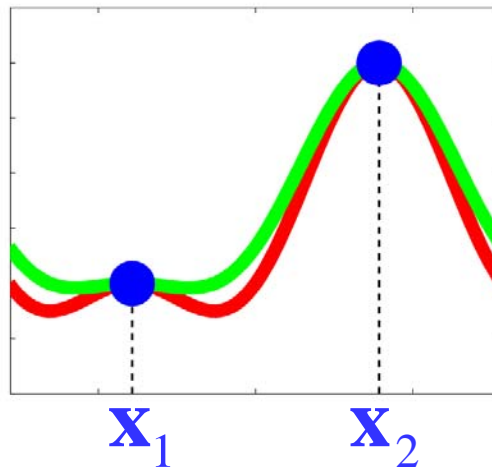
- **Learning methods:**

What is the best way to learn?

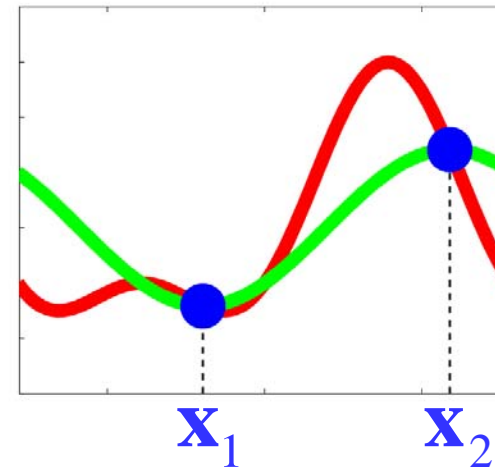
Active Learning

For obtaining good learning results, questions should be determined appropriately.

— Target function
— Learned function



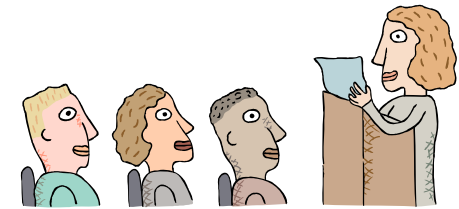
Good questions



Bad questions

Active Learning: Analogy to Real Life⁴⁰

- It is not interesting to **passively** attend the lecture.



- It is more effective to **actively** ask questions in the lecture.

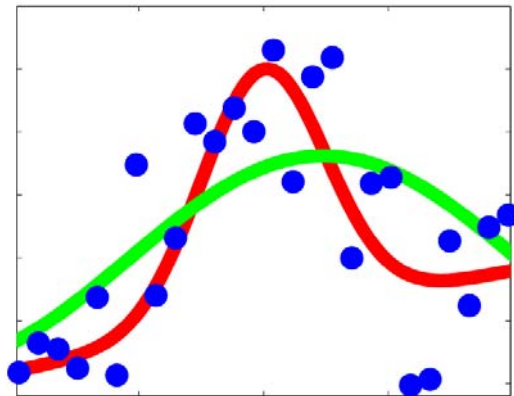


Model Selection

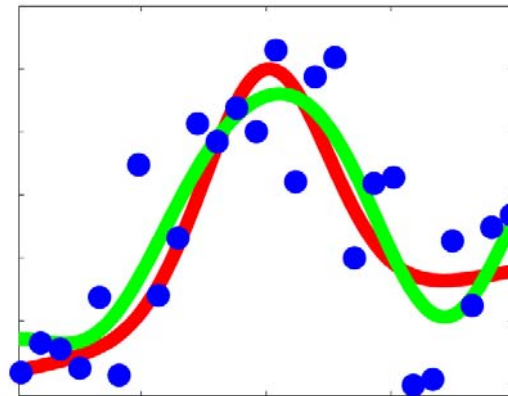
For obtaining good learning results, model should be determined appropriately.

Model is a set of functions from which learning results are searched.

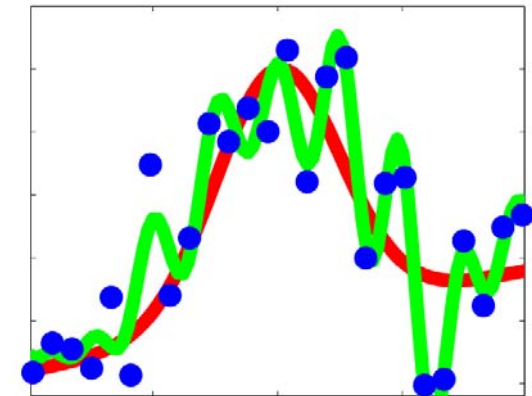
— Target function
— Learned function



Simple model



Appropriate model



Complex model

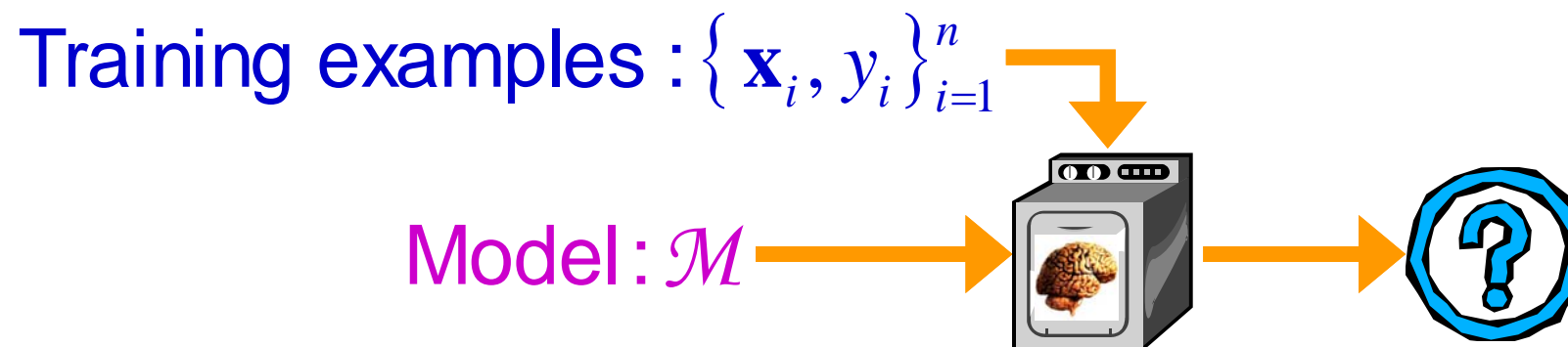
Model Selection: Analogy to Real Life

- A model represents **your ambition**.
- You learn a fixed amount of material.
- If you are **less ambitious**, you are not capable of even memorizing what you have learned. Therefore, you can not find the truth.
- If you are **too ambitious**, you can memorize what you have learned perfectly. However, you can not get the whole picture.
- If you are **appropriately ambitious**, then you can understand the truth.



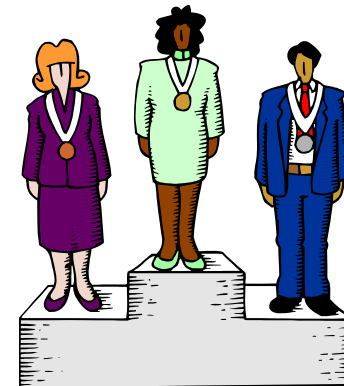
Learning Methods

- Now you have
 - **A model**, from which your learning result function is searched.
 - **Training examples**, which are pairs of questions and their answers.
- A learning method is **a rule to specify a function** in the model based on the training examples.



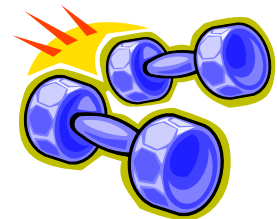
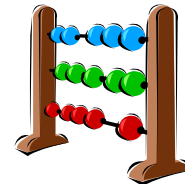
Learning Methods: Analogy to Real Life

- Now you have
 - Appropriate ambition for learning
 - Good questions and their answers
- What you should do is to just start studying!
- **Effectively** using your ambition and teaching materials is the key to success.



Conclusions

- There are 3 topics in learning research.
 - Understanding human brains
 - Developing learning machines
 - Mathematically clarifying mechanism of learning
- There are 3 types of learning.
 - Supervised learning
 - Unsupervised learning
 - Reinforcement learning
- Topics of supervised learning:
 - Active learning
 - Model selection
 - Learning methods



Homework

- Prepare your own supervised learning data sets (e.g., from **your research domain**)

$$\{(\mathbf{x}_i, y_i)\}_{i=1}^n$$

- Input \mathbf{x}_i should be vectors and output y_i should be scalars.

$$\mathbf{x}_i \in \mathbb{R}^d$$

$$y_i \in \mathbb{R}$$

- Better if

- Input is not so high dimensional
- The number of samples is large

$$d : \text{small}$$

$$n : \text{large}$$

- Explain specification of your data

Homework (cont.)

- Prepare a computer environment in which you can run
e.g., MATLAB, Octave, R, Scilab, ...
- Deadline: beginning of next class (April 17th)