# Introduction

Masashi Sugiyama
sugi@k.u-tokyo.ac.jp
http://www.ms.k.u-tokyo.ac.jp

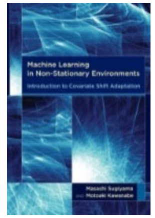
# About Myself

- **Masashi Sugiyama:**
  - Director: RIKEN AIP, Japan
  - Professor: University of Tokyo, Japan
  - Consultant: several local startups

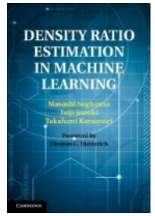
- **Interests: Machine learning (ML)**
  - ML theory & algorithm →
  - ML applications (signal, image, language, brain, robot, mobility, advertisement, biology, medicine, education…)

- **Academic activities:**
  - Program Chair: NeurIPS2015, AISTATS2019, ACML2010/2020, IEEE-CAI2024…
  - Keynote speaker: ICLR2023, IJCNN2024…

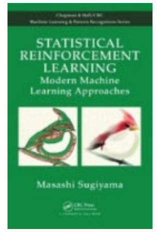
Sugiyama & Kawanabe, Machine Learning in Non-Stationary Environments, MIT Press, 2012



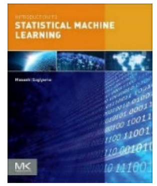
Sugiyama, Suzuki & Kanamori, Density Ratio Estimation in Machine Learning, Cambridge University Press, 2012



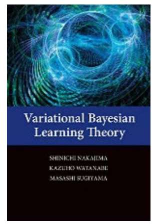

Sugiyama, Statistical Reinforcement Learning, Chapman and Hall/CRC, 2015



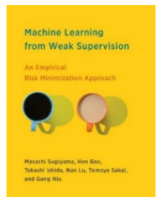

Sugiyama, Introduction to Statistical Machine Learning, Morgan Kaufmann, 2015




Nakajima, Watanabe & Sugiyama, Variational Bayesian Learning Theory, Cambridge University Press, 2019




Sugiyama, Bao, Ishida, Lu, Sakai & Niu. Machine Learning from Weak Supervision, MIT Press, 2022.

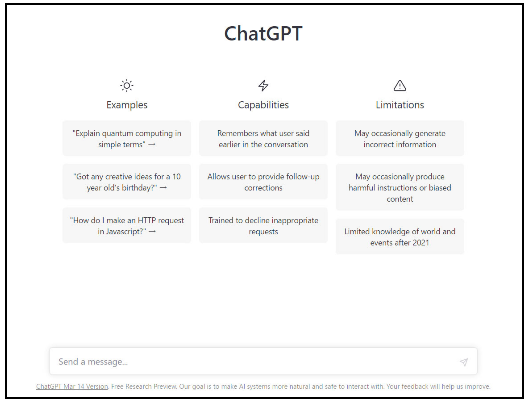


# Artificial Intelligence (AI)

- AI is used in various situations around us, such as self-driving cars, computer Go, and interactive bots.

http://www.cnbc.com/2015/10/14/tesla-rolls-out-autopilot-technology.html


http://gigazine.net/news/20160317-google-alphago/



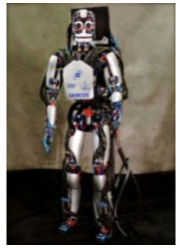

https://chat.openai.com/chat

- Behind these AI systems, machine learning techniques are used to make computers acquire human-like learning abilities.

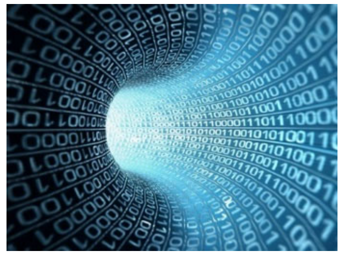
# Machine Learning (ML)

- Various usage of ML techniques:
  - Recognition of audio, images, and video
  - Information extraction from the web and SNS
  - Recommendation of products and services
  - Quality control of industrial products
  - Robotic system control
  - Medical image processing



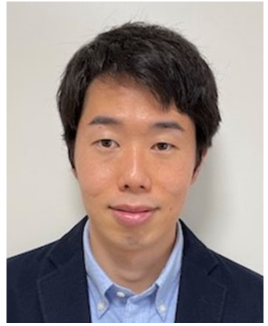



- With the advent of the big data era, ML technology is becoming increasingly important.




# Advanced Data Analysis

- Aim: We lecture fundamental concepts, mathematical foundations, and advanced algorithms of data analysis techniques:
  - Regression, classification, representation learning, clustering, outlier detection, density estimation, and deep learning.



Takashi Ishida

- Evaluation:
  - Regular reports (submit to UTOL before the next lecture).
  - Final report.

Masashi Sugiyama

- Lecture slides:

https://utol.ecc.u-tokyo.ac.jp/lms/course?idnumber=2025_4740_47130-53_01

# Lecture Schedule (2025S)

1. 04/8 Introduction
2. 04/15 Regression 1
3. 04/22 Regression 2

- 04/30 Cancelled

4. 05/13 Classification 1
5. 05/20 Classification 2
6. 05/27 Deep learning 1

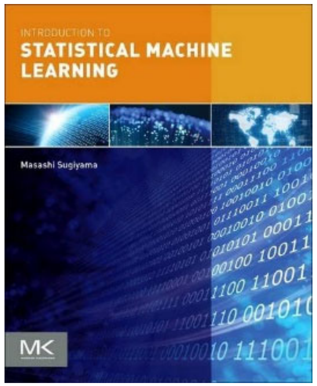
- 06/03 No lecture

7. 06/10 Deep learning 2
8. 06/17 Deep learning 3
9. 06/24 Semi-supervised learning
10. 07/01 Language models
11. 07/08 Representation learning 1
12. 07/15 Representation learning 2
13. 07/22 Advanced topics

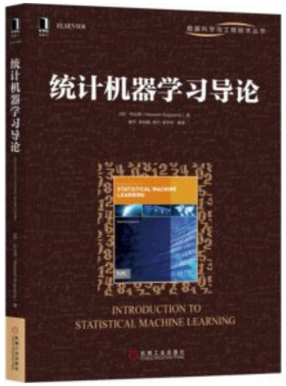
- Zoom URL:
  https://u-tokyo-ac-jp.zoom.us/j/82403666641?pwd=KRS6iPnUXrLn3Aa79Qe3fdybxADwkS.1
- UTOL: https://utol.ecc.u-tokyo.ac.jp/lms/course?idnumber=2025_4740_47130-53_01

# Textbook(s)

- M. Sugiyama: Introduction to Statistical Machine Learning, Morgan Kaufmann, 2015.
  - 杉山：イラストで学ぶ機械学習，講談社，2013
  - 杉山：機械学習のための確率と統計，講談社，2015.
  - 杉山：图解机器学习, 2015
  - 杉山：统计机器学习导论, 2018
  - 스기야마: 그림과 수식으로 배우는 통통 머신러닝, 2017
  - 2nd edition under preparation (with T. Ishida).



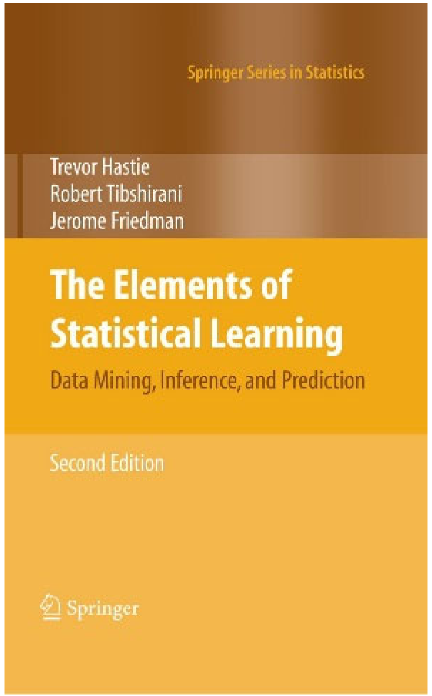


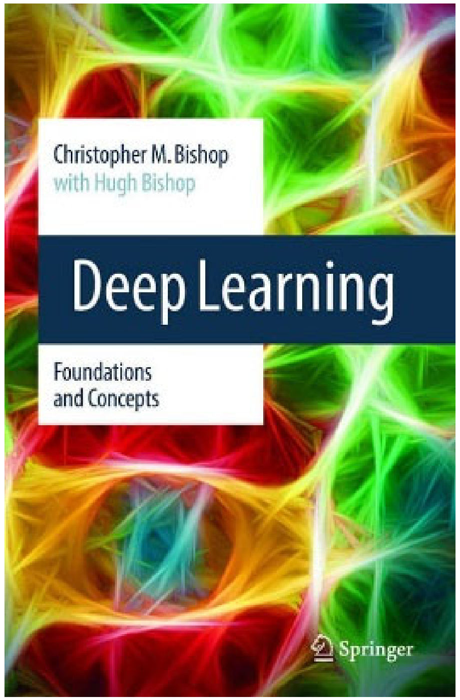




# Reference Books

- T. Hastie, R. Tibshirani & J. Friedman: The Elements of Statistical Learning, 2nd edition, Springer, 2009

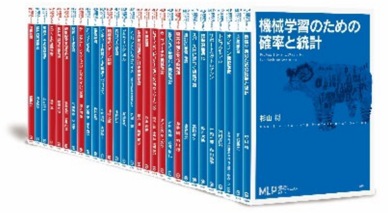
- C. M. Bishop, Deep Learning, Springer, 2024. https://www.bishopbook.com/



# その他日本語の参考書






## 講談社：機械学習プロフェッショナルシリーズ

- 機械学習のための連続最適化：
  金森敬文，鈴木大慈，竹内一郎，佐藤一誠
- オンライン予測：畑埜晃平，瀧本英二
- 関係データ学習：石黒 勝彦，林 浩平
- データ解析におけるプライバシー保護：佐久間　淳
- ウェブデータの機械学習：
  ダヌシカ ボレガラ，岡﨑直観，前原貴憲
- バンディット問題の理論とアルゴリズム：
  本多淳也，中村篤祥
- グラフィカルモデル：渡辺有祐
- ヒューマンコンピュテーションとクラウドソーシング：
  鹿島久嗣，小山聡，馬場雪乃
- ノンパラメトリックベイズ　点過程と統計的機械
  学習の数理：佐藤一誠
- 変分ベイズ学習：中島伸一
- スパース性に基づく機械学習：冨岡亮太
- 生命情報処理における機械学習多重検定と
  推定量設計：瀬々潤，浜田道昭
- 劣モジュラ最適化と機械学習：河原吉伸，永野清仁
- 統計的学習理論：金森敬文
- 確率的最適化：鈴木大慈
- 異常検知と変化検知：井手剛，杉山将
- サポートベクトルマシン：竹内一郎，烏山昌幸
- 機械学習のための確率と統計：杉山将
- 深層学習：岡谷貴之
- オンライン機械学習：
  海野裕也，岡野原大輔，得居誠也，徳永拓之
- トピックモデル：岩田具治
- 統計的因果探索：清水昌平
- 画像認識：原田達也
- 深層学習による自然言語処理：坪井祐太，
  海野裕也，鈴木潤
- 音声認識：篠田浩一
- ガウス過程と機械学習：持橋大地，大羽成征
- 強化学習：森村哲郎
- ベイズ深層学習：須山敦志
- 機械学習工学：石川冬樹，丸山宏（編著）
- 最適輸送の理論とアルゴリズム：佐藤竜馬
- 更に新刊準備中！

# Contents

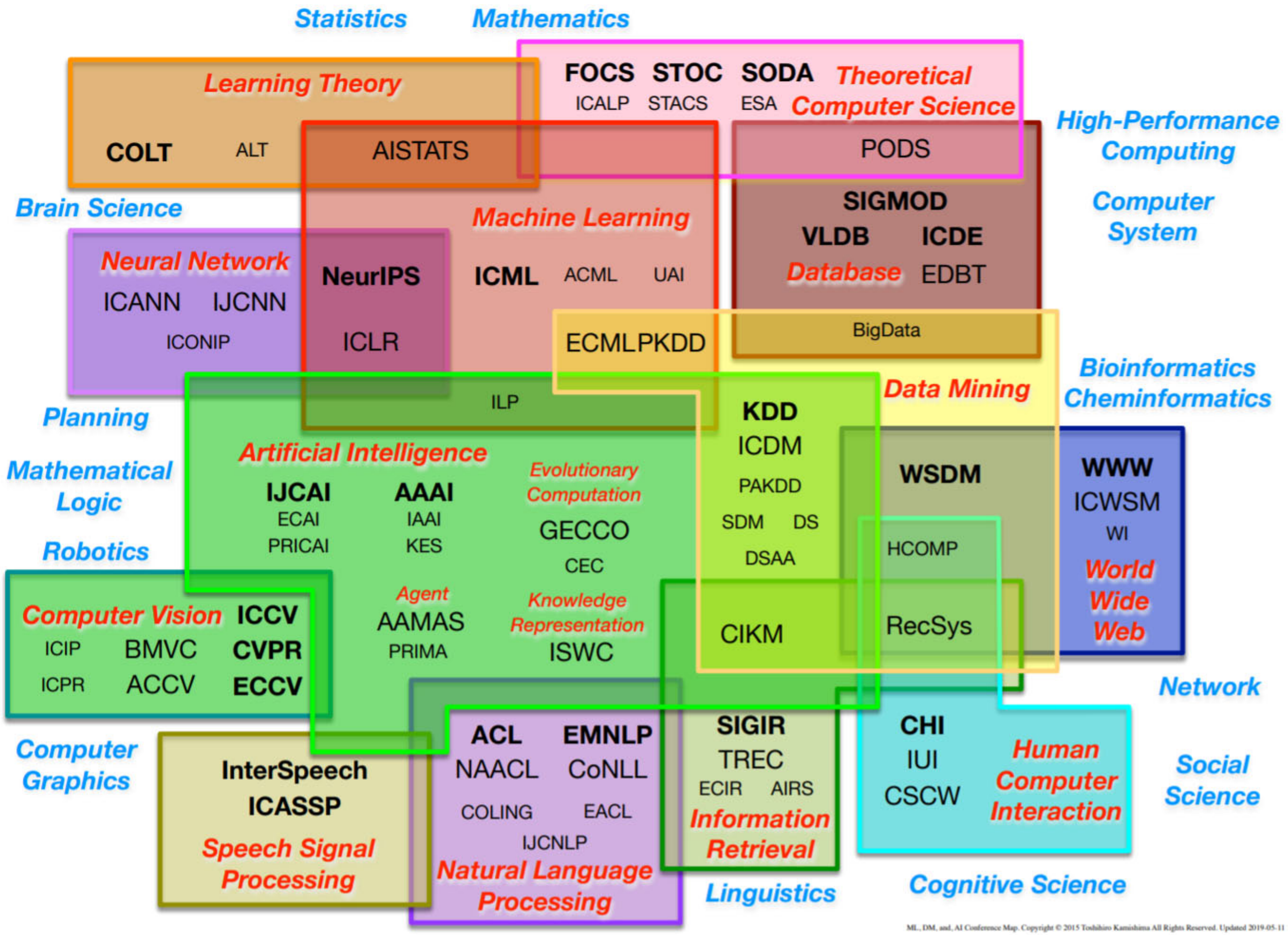


1. Trends of ML/AI Research
   A) International Conferences
   B) RIKEN-AIP
2. Basics of Machine Learning
3. Basics of Supervised Learning
4. Summary

# AI-related Research Communities



- Researchers don't publish papers in journals but international conferences (ranging from fundamental math to real-world application).

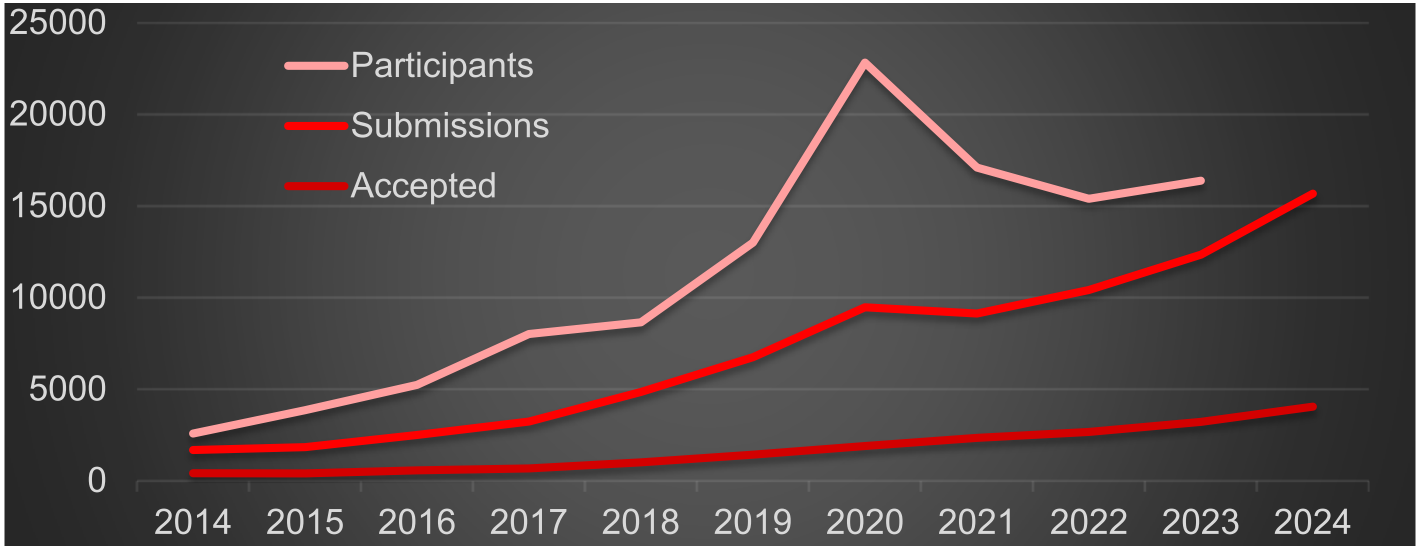



http://www.kamishima.net/archive/MLDMAImap.pdf

# Trends in NeurIPS

NEURAL INFORMATION PROCESSING SYSTEMS

- Sharp increase in participants and papers.




- Industry sponsorship is also very active:
  - Early 2000s: US-based IT giants
  - Late 2000s: IT giants around the world
  - 2010s: High diversity (IT and beyond, start ups to giants)
  - 2020s: Shrunk sponsorship

# Trend Transition



- Around 2015:
  - Discussion focused on ML technology itself.
  - Expectations for further development of the technology were rising with the appearance of AlphaGo, self-driving cars, conversational robots, etc.
  - North American companies and universities dominated both research and business
- Now:
  - Competition to develop ML technology intensifies further.
  - Application of ML to other scientific research areas.
  - Addressing social issues such as equity and explainability.
  - Emphasis on diversity, including support for minorities: Women in ML, Black in AI, Queer in AI, Global South in AI,…

# Nikkei Jan. 10th, 2025



Number of authors in NeurIPS2024, ICML2024, and ICLR2024.

| Rank '24 | ('20) | Organization | Authors |
|---|---|---|---|
| 1 | (1) | Google (US) | 2174 |
| 2 | (8) | Tsinghua U (CH) | 1748 |
| 3 | (2) | Stanford U (US) | 1081 |
| 4 | (3) | MIT (US) | 1037 |
| 5 | (7) | CMU (US) | 1015 |
| 6 | (17) | Peking U (C) | 906 |
| 6 | (89) | Zhejiang U (C) | 906 |
| 8 | (5) | Microsoft (US) | 851 |
| 9 | (40) | Shanghai Jiao Tong U (C) | 810 |
| 10 | (6) | Meta (US) | 782 |
| ⋮ | | | ⋮ |
| 64 | (44) | RIKEN (JP) | 188 |
| ⋮ | | | ⋮ |
| 71 | (50) | U Tokyo (JP) | 171 |

- US universities and companies and Chinese universities are top-ranked.
- Universities in Singapore, UK, Korea, Switzerland, and Canada follow.
- RIKEN is more than doubled compared with 2020, but its ranking is lowered.

https://www.nikkei.com/article/DGXZQOSG108LH0Q4A211C2000000/

# Contents

# RIKEN's Organizational Structure

2,770 researchers
670 admin staffs
(as of Apr. 2024)

**TRIP Initiative**

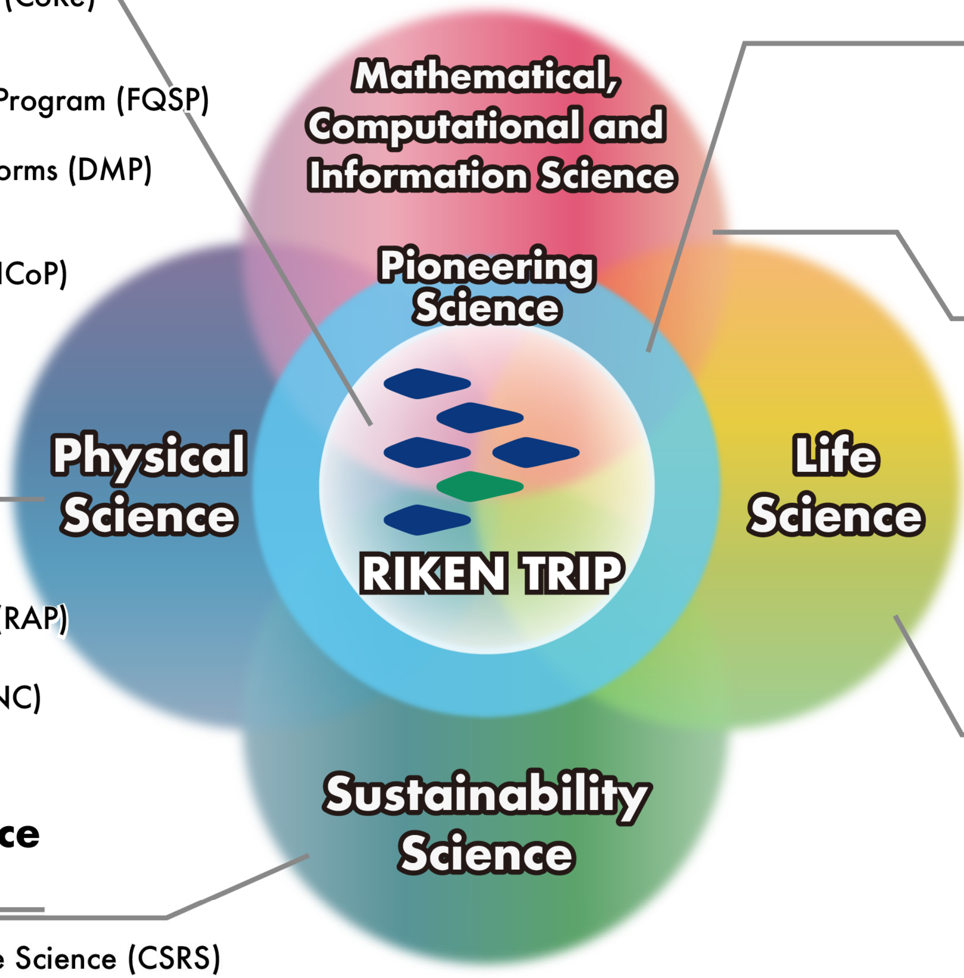
Data and Computational Sciences Integration Research Program (CoRe)
Advanced General Intelligence for Science Program (AGIS)
Fundamental Quantum Science Program (FQSP)
Program for Drug Discovery and Medical Technology Platforms (DMP)
Advanced Semiconductor Science Program (ASSP)
Industrial Co-creation Program (ICoP)
Baton Zone Program (BZP)

**Physical Science Domain**

Center for Emergent Matter Science (CEMS)
Center for Advanced Photonics (RAP)
Nishina Center for Accelerator-Based Science (RNC)
SPring-8 Center (RSC)

**Sustainablility Science Domain**

Center for Sustainable Resource Science (CSRS)
(BioResource Research Center (BRC))




**Pioneering Science Domain**

Pioneering Research Institute (PRI)

**Mathematical, Computational and Information Science Domain**

Center for Interdisciplinary Theoretical and Mathematical Sciences (iTHEMS)
Center for Computational Science (R-CCS)
Center for Quantum Computing (RQC)
Center for Advanced Intelligence Project (AIP)
RIKEN Information R&D and Strategy Headquarters (R-IH)

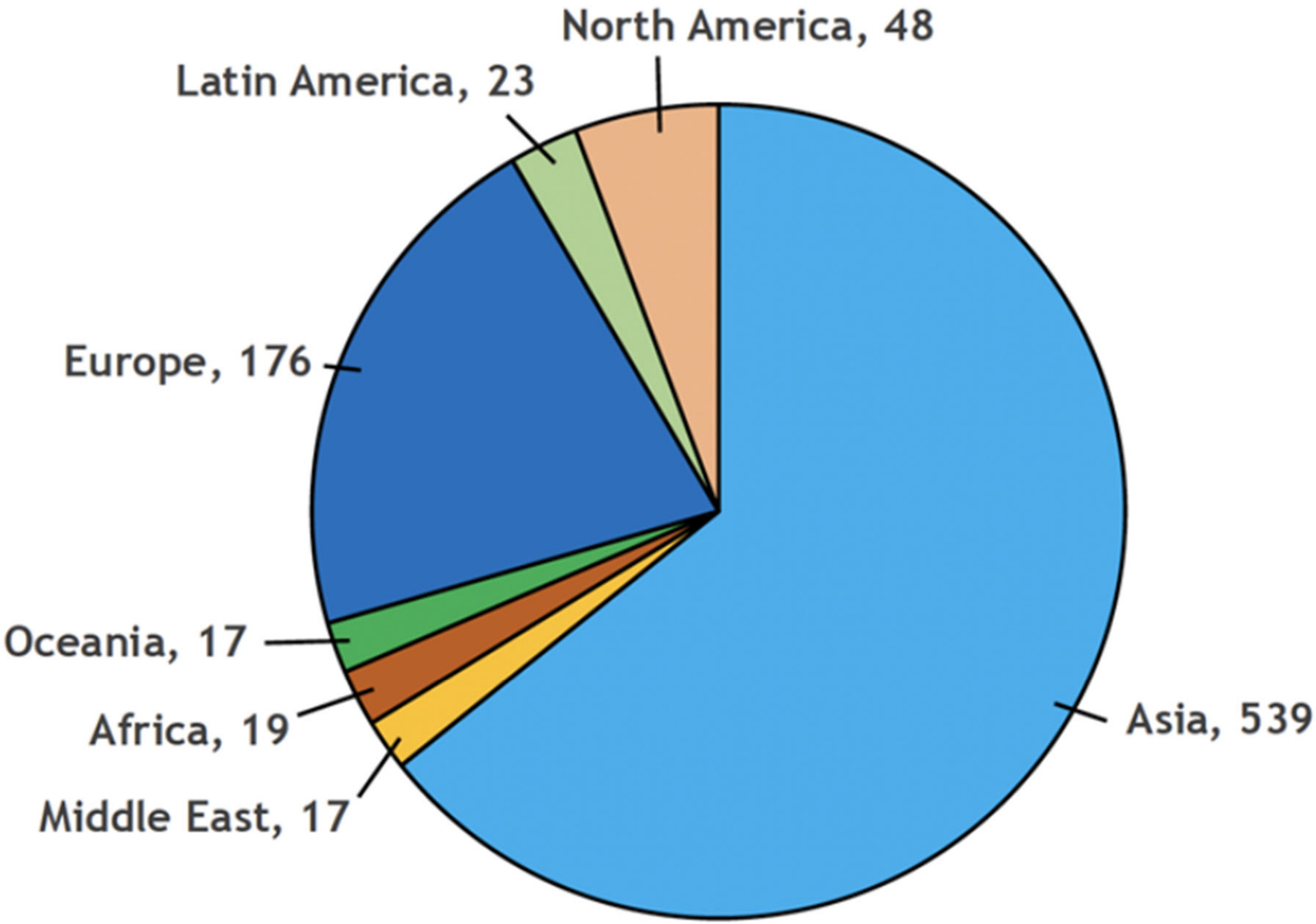
**Life Science Domain**

Center for Integrative Medical Sciences (IMS)
Center for Biosystems Dynamics Research (BDR)
Center for Brain Science (CBS)
BioResource Research Center (BRC)

*TRIP (Transformative Research Innovation Platform of RIKEN platforms)

# 839 International Researchers






Breakdown of RIKEN scientists by country/region** (as of Oct 1, 2024)

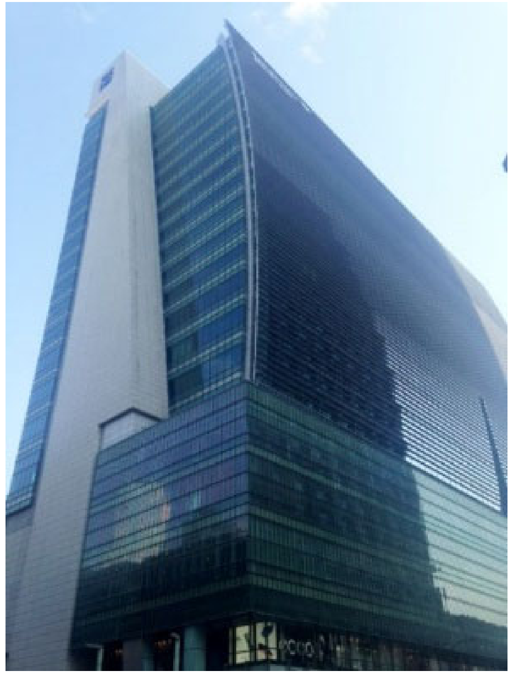
# What is RIKEN-AIP?



- RIKEN founded Center for Advanced Intelligence Project (AIP) in 2016, under Ministry of Education, Culture, Sports, Science and Technology (MEXT).

MEXT
MINISTRY OF EDUCATION, CULTURE, SPORTS, SCIENCE AND TECHNOLOGY-JAPAN

Main office located in the heart of Tokyo



Distributed office across Japan


In-house GPU servers

Open discussion space

# 3 Research Groups



## Generic Technology Research Group (16)

- Imperfect Information Learning Team

Masashi Sugiyama
- Structured Learning Team

Yoshinobu Kawahara
- Tensor Learning Team

Qibin Zhao
- Functional Analytic Learning Team

Quang Minh Ha
- Succinct Information Processing Team

Yasuo Tabei
- Deep Learning Theory Team

Taiji Suzuki
- Computational Learning Theory Team

Kohei Hatano
- Causal Inference Team

Shohei Shimizu
- Approximate Bayesian Inference Team

Mohammad Emtiyaz Khan
- Continuous Optimization Team

Akiko Takeda
- Mathematical Science Team

Kenichi Bannai
- High-Dimensional Causal Analysis Team

Masaaki Imaizumi
- Sequential Decision Making Team

Shinji Ito
- Uncertainty Quantification Team
Futoshi Futami
- Information Statistical Mechanics and Dynamics Team
Ayaka Sakata
- Computational Social Choice Team
Ayumi Igarashi

## Goal-Oriented Technology Research Group (25)

- Cancer Translational Research Team

Ryuji Hamamoto
- Medical-risk Avoidance based on iPS Cells Team

Naonori Ueda
- Molecular Informatics Team

Koji Tsuda
- Cognitive Behavioral Assistive Technology Team

Mihoko Otake
- Disaster Resilience Science Team

Naonori Ueda
- Robotics for Infrastructure Management Team

Takayuki Okatani
- Machine Intelligence for Medical Engineering Team

Tatsuya Harada
- Data-Driven Biomedical Science Team

Ichiro Takeuchi
- Computational Brain Dynamics Team

Okito Yamashita
- Statistical Genetics Team

Gen Tamiya
- Pathology Informatics Team

Yoichiro Yamamoto
- Natural Language Understanding Team

Kentaro Inui
- Knowledge Acquisition Team

Yuji Matsumoto
- Geoinformatics Team

Naoto Yokoya
- Music Information Intelligence Team

Masatoshi Hamanaka
- Sound Scene Understanding Team

Kazuyoshi Yoshii
- Robot Learning Team

Takayuki Osa
- Computational Physical Machine Learning Team
Takaharu Yaguchi
- AI Computing Team
Shinya Takamaeda
- 3D Environmental Information Understanding Team
Asako Kanezaki
- Robot System Team
Kei Okada
- (Under adjustment)
- Explainable AI Team
Hitomi Yanaka

## AI in Society Research Group (6)

- AI Utilization in Society and Legal System Team

Hiroshi Nakagawa
- Science, Technology and Society Team

Osamu Sakura
- Decentralized Big Data Team

Koiti Hasida
- Business and Economic Information Fusion Analysis Team

Takahiro Hoshino
- AI Security and Privacy Team

Jun Sakuma
- AI Safety and Reliability Unit

Hiromi Arai

---

- Learning & inference algorithm:
  - Robust, Bayes, high-dim…
- Statistical & optimization theory:
  - Convergence, non-convex, …

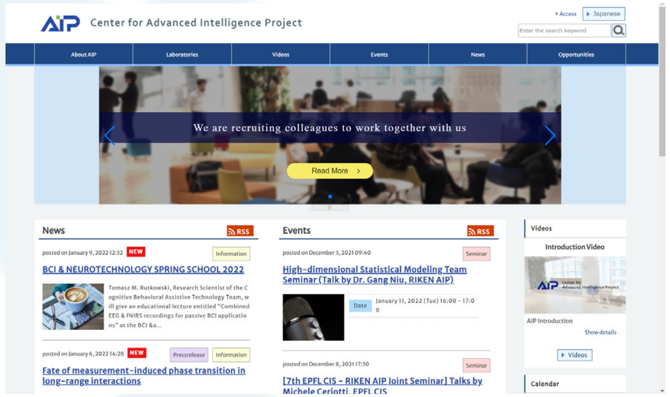
- AI for science:
  - Cancer, material, genomics…
- AI for societal problems:
  - Natural disaster, elderly healthcare, education…

- Personal data management
- AI security
- Economy and management
- Ethics and governance

# Statistics

## Diverse research staffs:

- 142 employed researchers (44% international, 22% female)
- 293 visiting researchers
- 120 domestic students
- 190 international interns (total)





As of 2024 Dec.

# Lecture by RIKEN-AIP Researchers



Special Topics in Mechano-Informatics II (in English, online)
Wednesday 14:55-16:40
知能機械情報学特別講義II （英語・オンライン）
https://utol.ecc.u-tokyo.ac.jp/lms/course?idnumber=2025_4885_4850-1022_01

1. 4/9 Masashi Sugiyama,
   Overview of AI Research and Introduction of RIKEN-AIP
2. 4/16 Jingfeng Zhang,
   Trustworthy Machine Learning
3. 4/23 Lin Gu,
   Latest Advancement of Medical AI

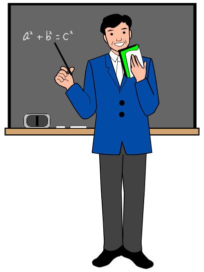
- 4/30 No class

4. 5/7 Nan Lu,
   Machine Learning from Imperfect Data
5. 5/14 Qibin Zhao,
   Interpretable and Adversarial Machine Learning
6. 5/21 Chao Li,
   Advancing Machine Learning with Tensor Networks
7. 5/28 Tatsuya Harada,
   Visual Recognition
8. 6/4, Minh Ha Quang,
   Geometrical Methods in Machine Learning
9. 6/11 Okan Koc,
   Introduction to Domain Adaptation
10. 6/18 Okan Koc,
    Introduction to Reinforcement Learning
11. 6/25 Yasuo Tabei,
    Succinct Data Structure for Scalable Knowledge Discoveries
12. 7/2 Tomasz Rutkowski,
    AI and Human Brain: Machine Learning Applications in Brain-Computer Interface
13. 7/9 Tomasz Rutkowski,
    AI for Time-Series Analysis: Wellbeing and Multimedia Applications

# Contents



1. Trends of ML/AI Research
2. Basics of Machine Learning
   A) Supervised Learning
   B) Unsupervised Learning
   C) Reinforcement Learning
3. Basics of Supervised Learning
4. Summary

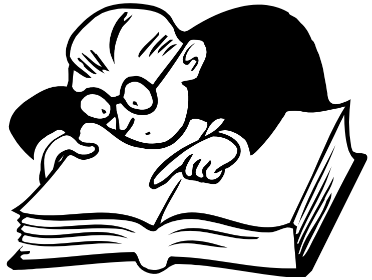


# 3 Types of Machine Learning

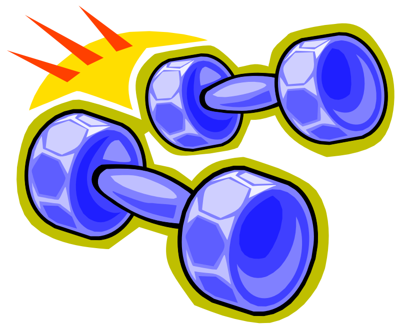


- **Supervised learning:**
  Human helps computer.

- **Unsupervised learning:**
  Computer learns itself.



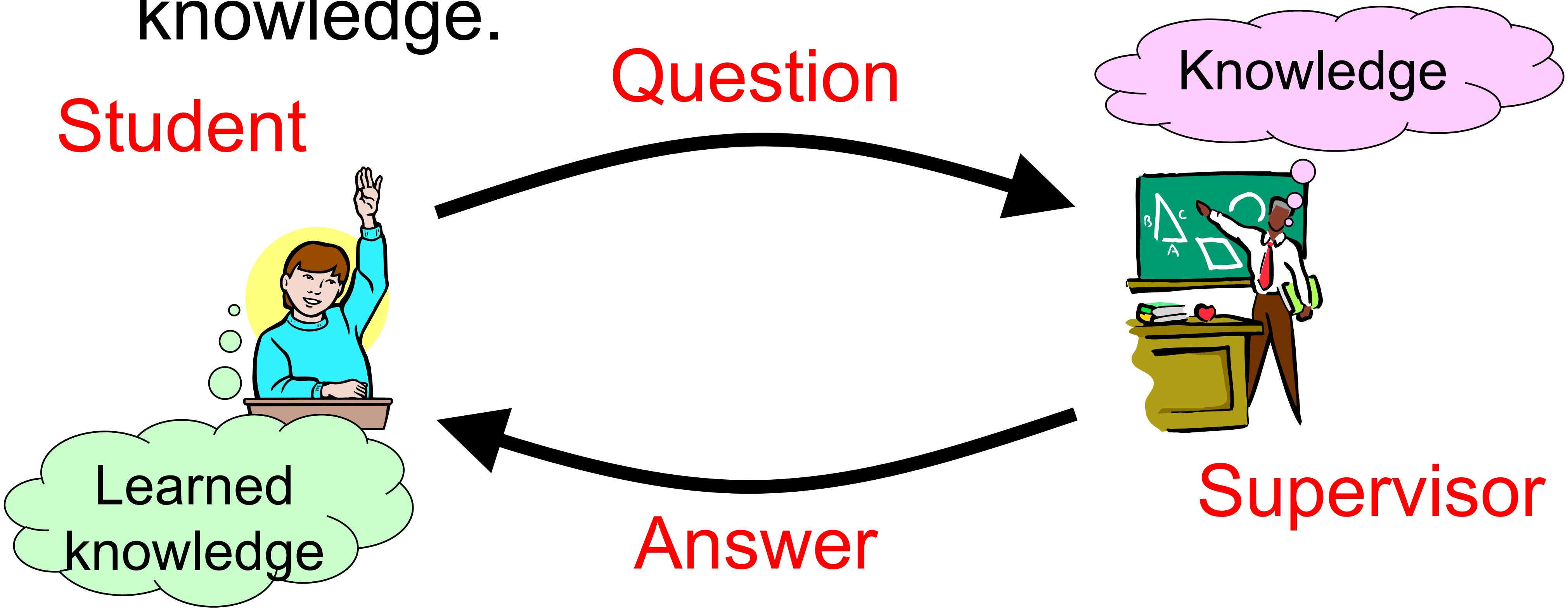
- **Reinforcement learning:**
  Computer agent learns through interaction with environment.



# Contents

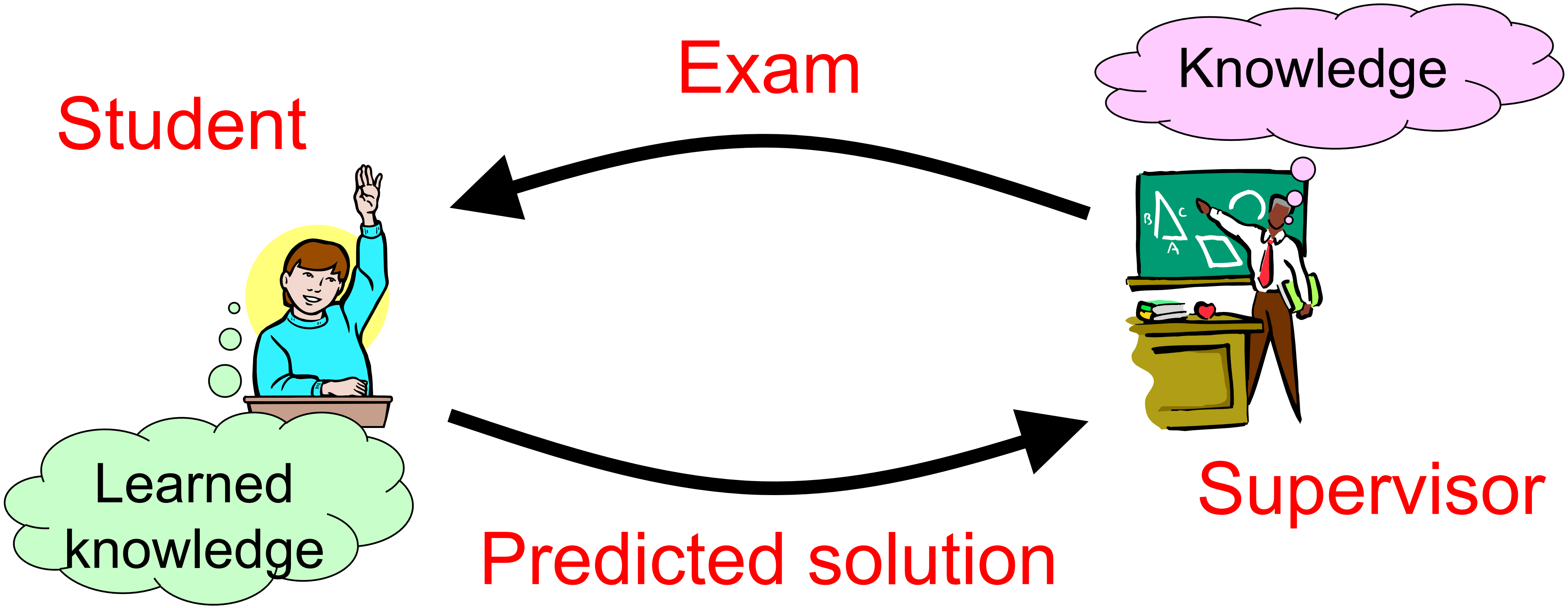
# Supervised Learning

- Supervisor: Can answer any question.
- Student: Ask a question to the supervisor and receive an answer.
- Goal: Let the student learn the supervisor's knowledge.



# Generalization Ability

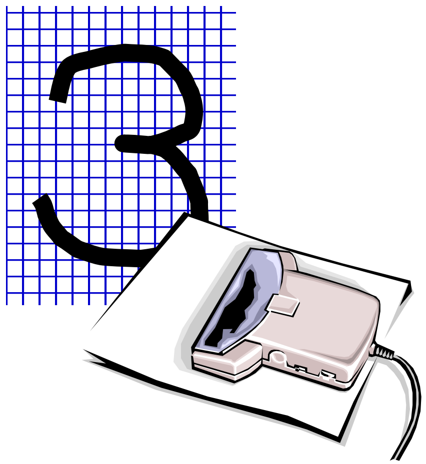
- If the student learns all of the teacher's knowledge instead of just memorizing the taught answers, the student can correctly answer new questions that he or she has never seen before.



# Ex: Hand-Written Digits

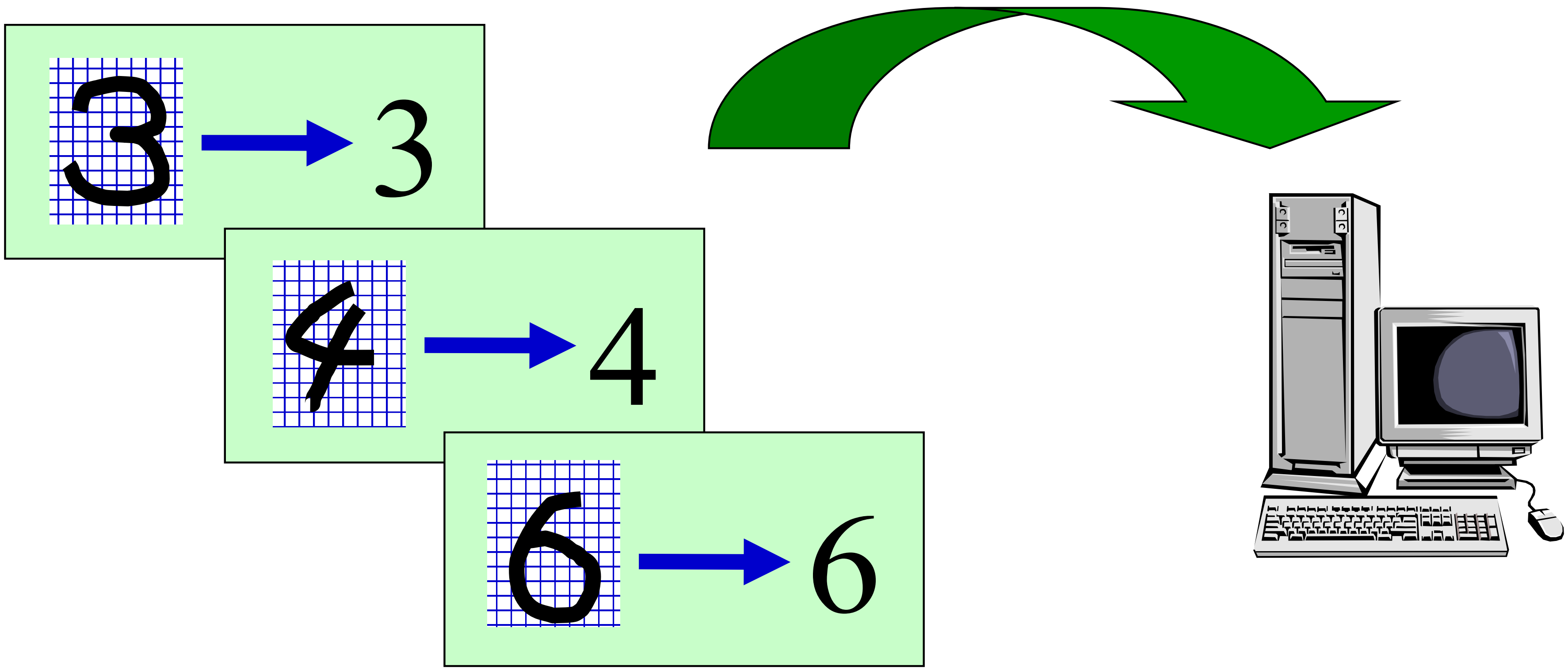
- Digit recognition:

  Let the computer to recognize scanned digits.



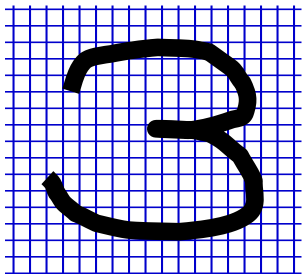
# Ex: Hand-Written Digits

- Given an image of digit, the supervisor should be able to answer what the written number is (we can do!).
- Let the computer learn many examples.



# Ex: Hand-Written Digits

- Recent computers have huge memory space. Can they just memorize all digit images?
- Suppose a digit image consists of 10x10 pixels and each pixel has 256 gray-level.



# Ex: Hand-Written Digits

- The number of all possible images is

$$256^{100} = (2^8)^{100} = (2^{10})^{80} \approx (10^3)^{80} = 10^{240}$$

- It is way impossible to memorize all.
- It is even not possible to prepare such data manually!

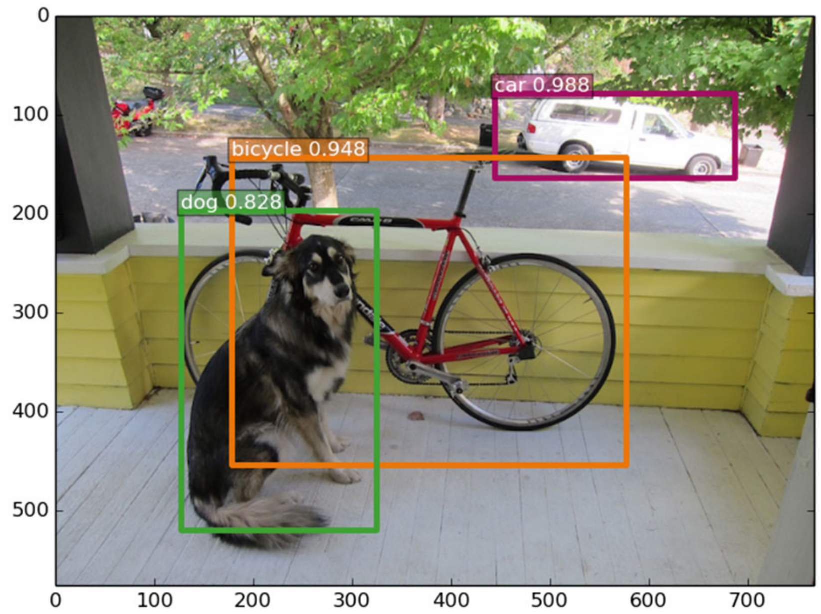
# More Examples



Speech recognition
(speech to language)

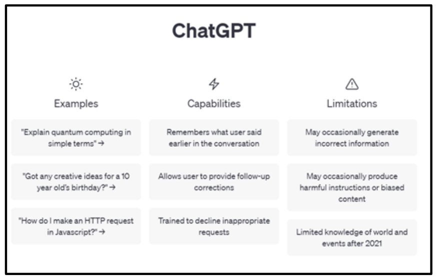
https://spjai.com/speech-recognition/

Image understanding
(image to objects)

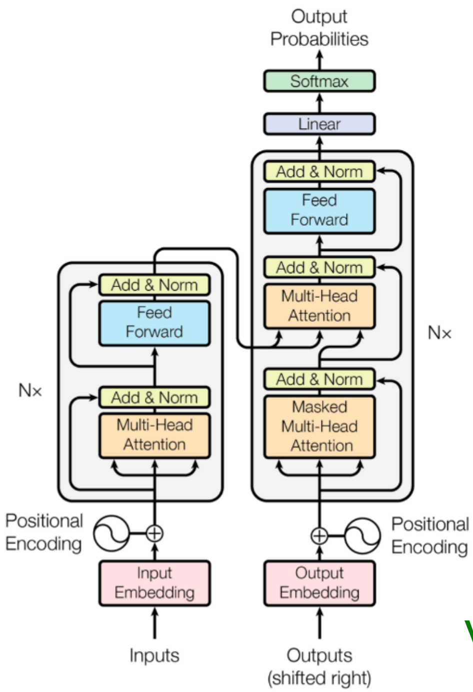



https://github.com/zhreshold/mxnet-ssd

Chatbot
(word sequence
to the next word)




https://chat.openai.com/



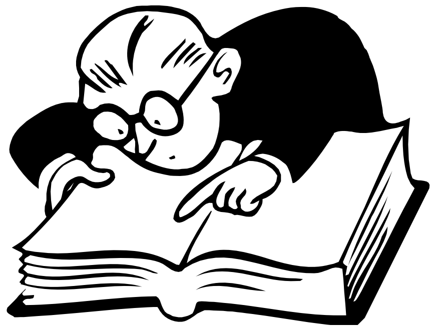

Vaswani+ (2017)

# Contents

# Unsupervised Learning

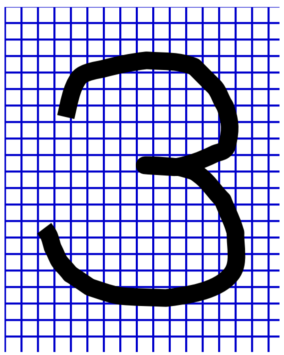


- No supervisor exists.
  - Computer learns by itself.
- Goal: Find "useful" knowledge from data.
  - But the "usefulness" needs to be specified by us.
- Sometimes called "data mining".



# Data Visualization

- Suppose a digit image consists of 10x10 pixels:
  - It is a 100-dimensional vector.

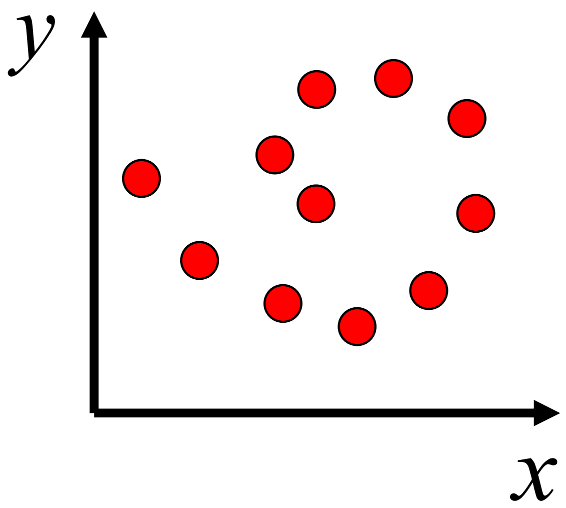


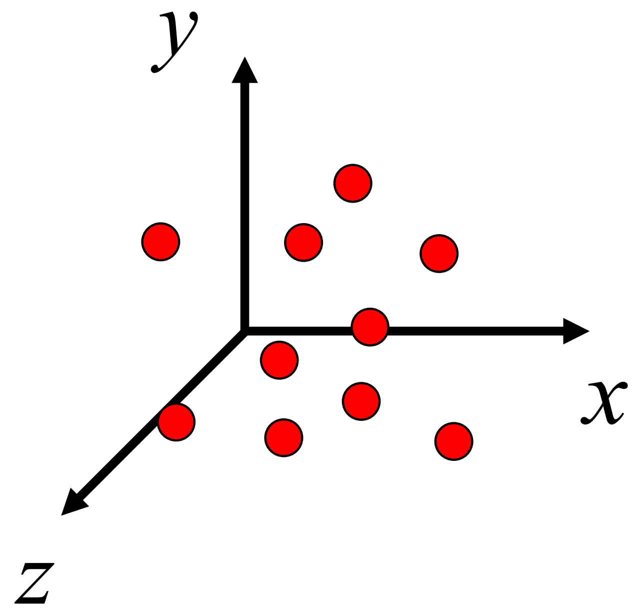
$$\left.\begin{pmatrix} 3 \\ 5 \\ 2 \\ \vdots \\ 4 \end{pmatrix}\right\} \begin{array}{c}\text{100} \\ \text{dimensions}\end{array}$$

# Data Visualization

- 2d or 3d data can be easily visualized.
  - We can obtain geometric intuition!

$\begin{pmatrix} x \\ y \end{pmatrix}$



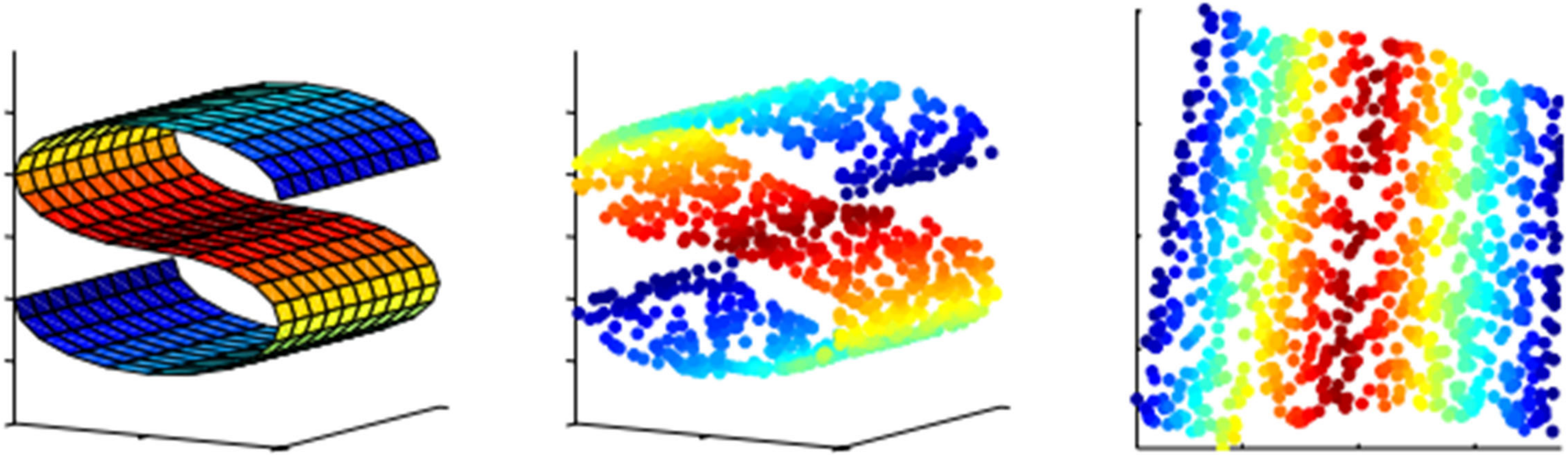

$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$

- However, 100d data cannot be visualized!

# Data Visualization

- Goal: Reduce the data dimensionality (to 2 or 3 usually), while keeping intrinsic properties of the original data.

2d

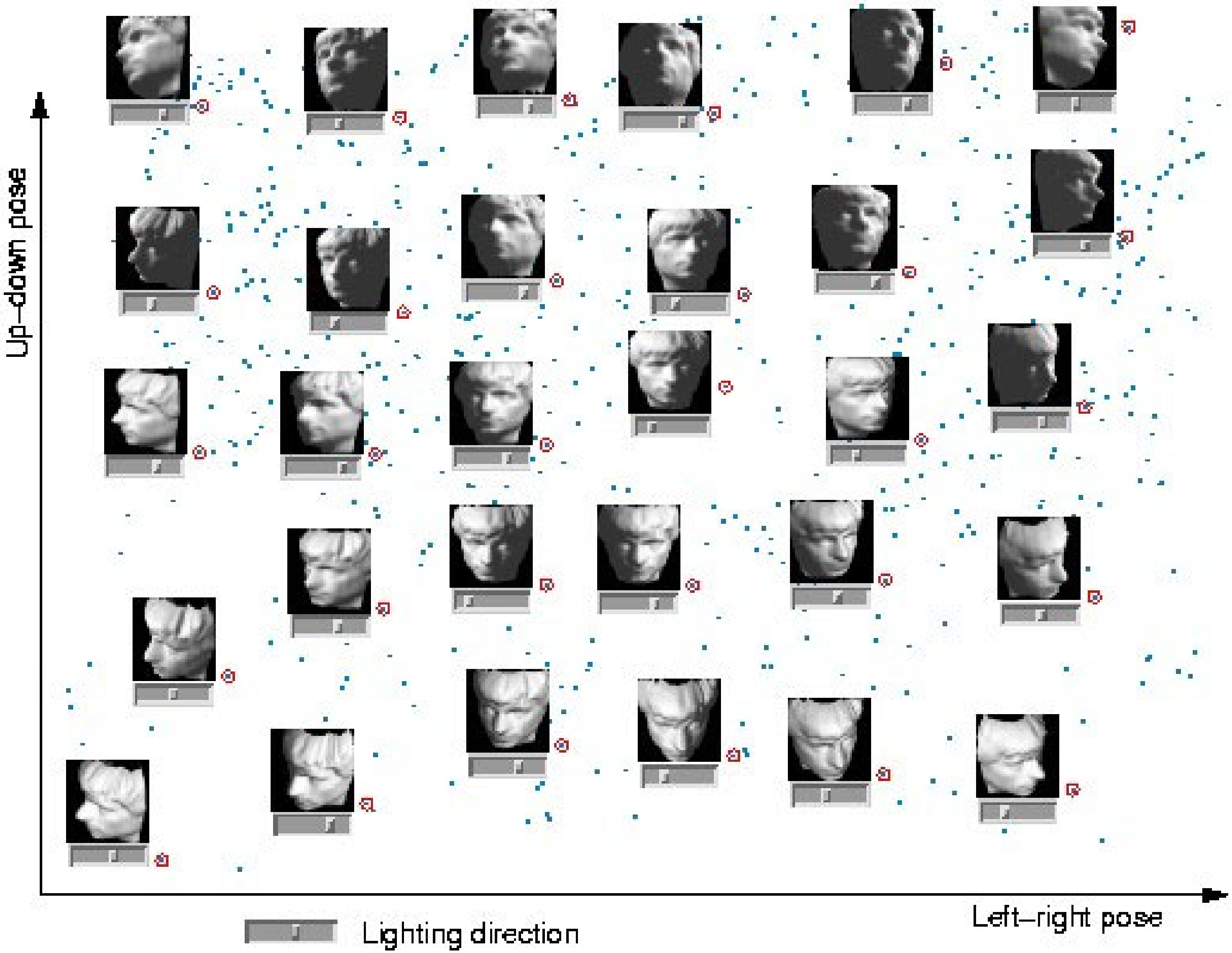


3d 3d

http://www.jmlr.org/papers/volume4/saul03a/saul03a.pdf

# Example

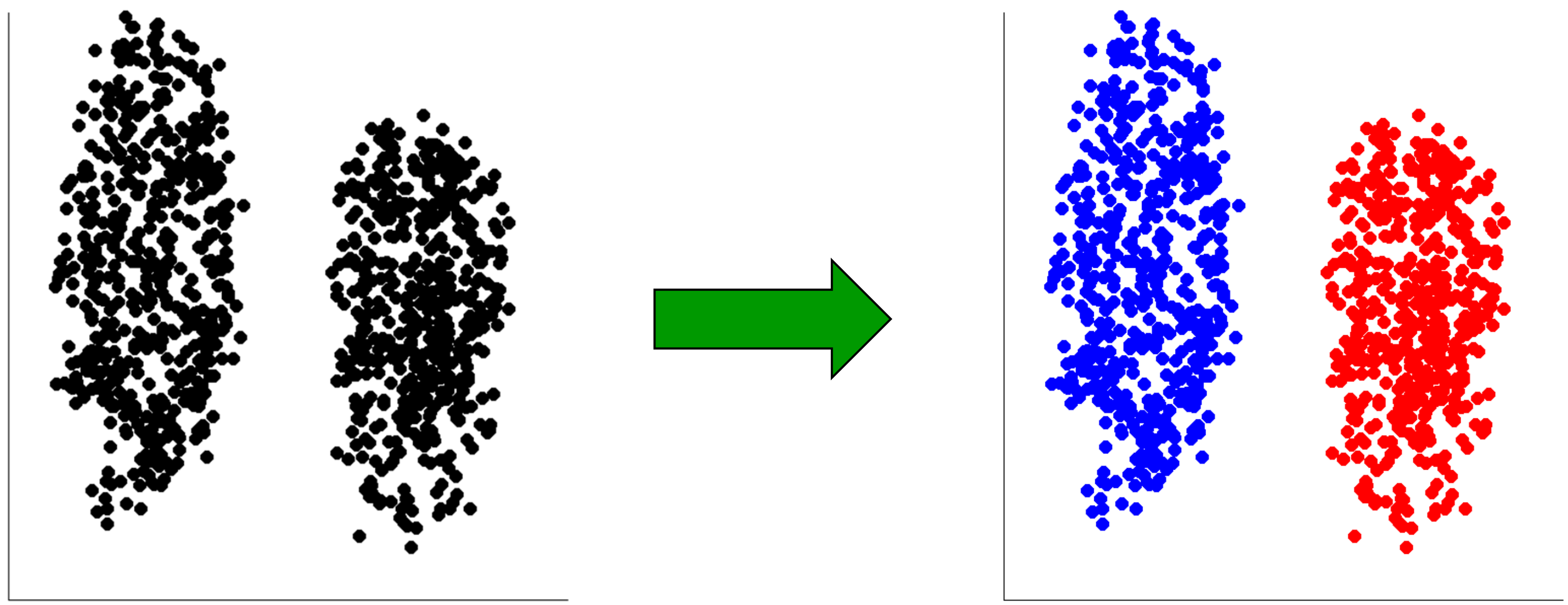
- Distribution of face images (64x64=4096):



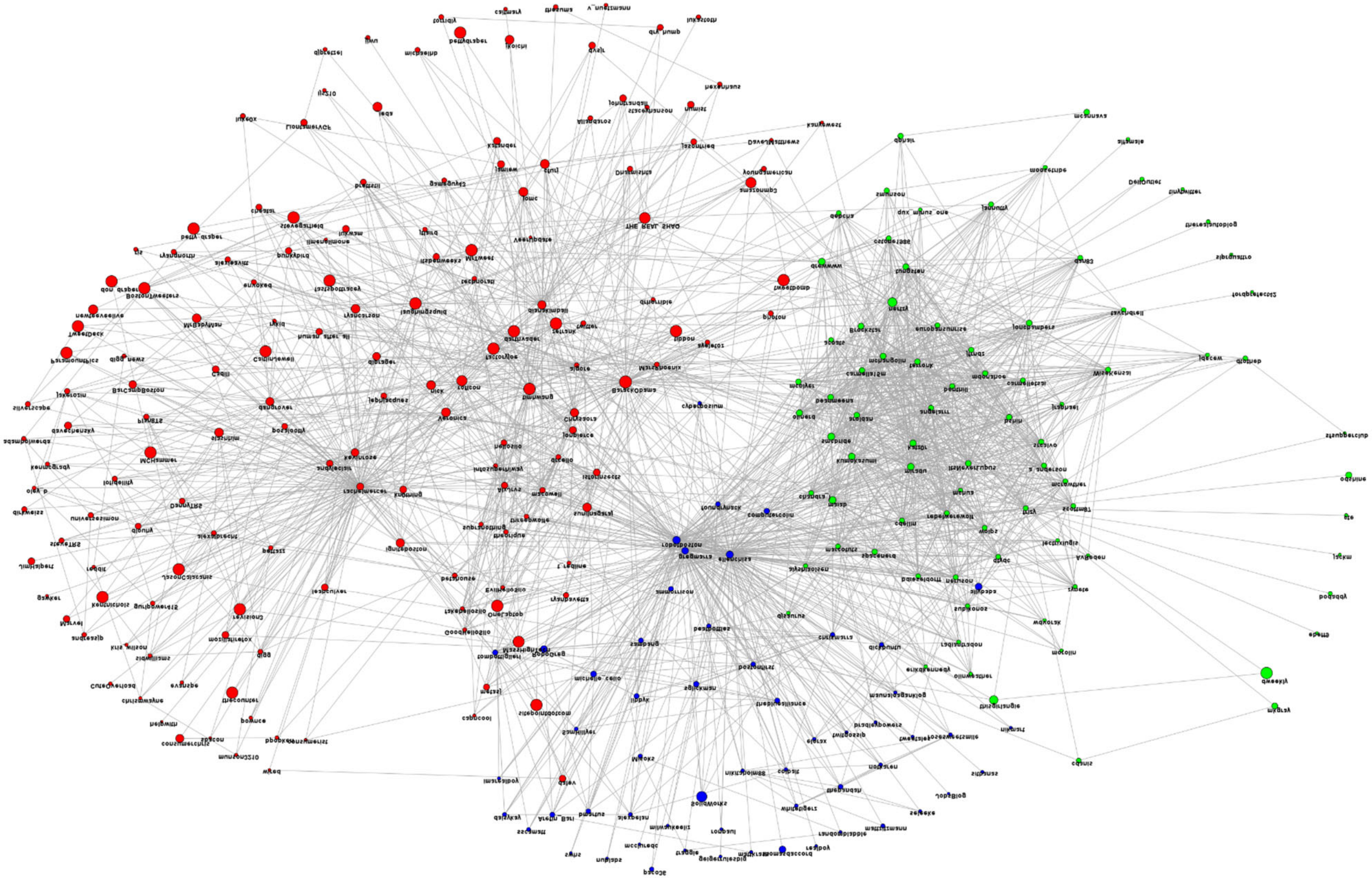

http://isomap.stanford.edu/

# Data Clustering

- Goal: Group given data so that
  - Similar points belong to the same cluster;
  - Dissimilar points belong to different clusters.



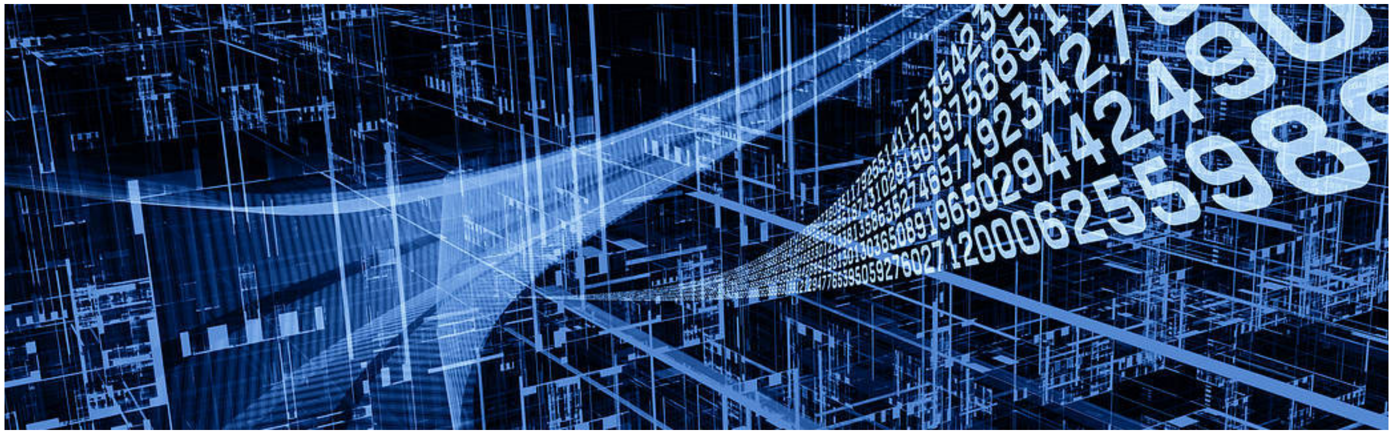
# Example: Community Discovery in Twitter



http://ca.olin.edu/2008/realboy/images/world.png

# Anomaly Detection

- **Goal:** We want to know the system anomalies.
  - Issue an alert if data different from the norm appear.
- **Examples:**
  - Intrusion detection in computer networks.
  - Anomaly detection in satellites.

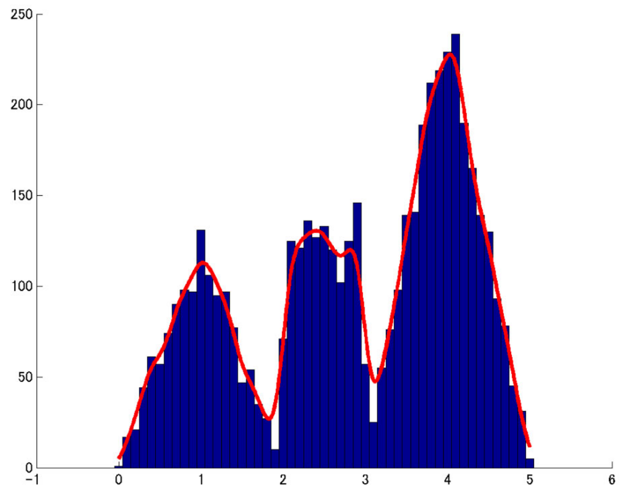


https://www.nasa.gov/digitalstrategy

http://marvin.news/2324

# Density Estimation

- Goal: Given data, estimate their probability density function.
  - Once the probability density is estimated, new data can be generated!




Ho+ (2020)

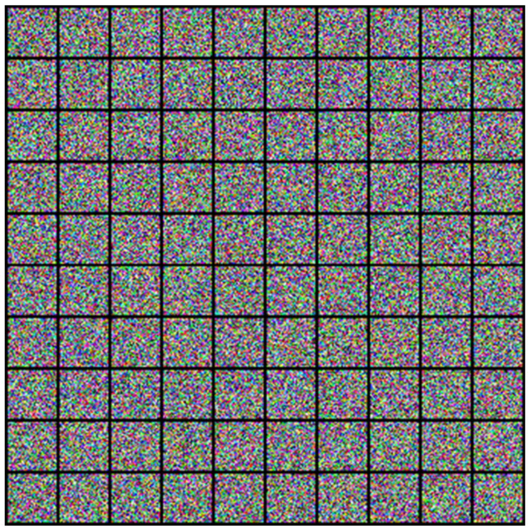
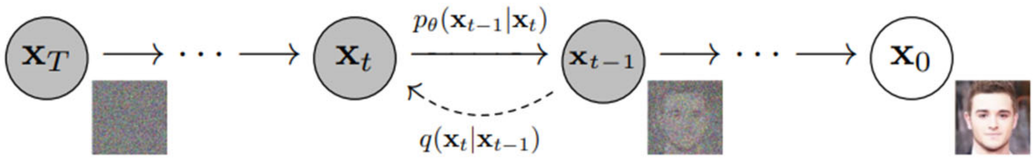


Image generation

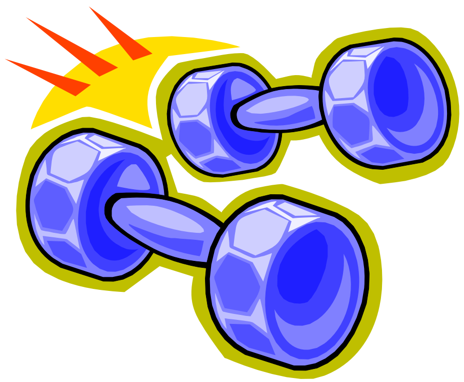



https://yang-song.net/blog/2021/score/

# Contents

# Reinforcement Learning

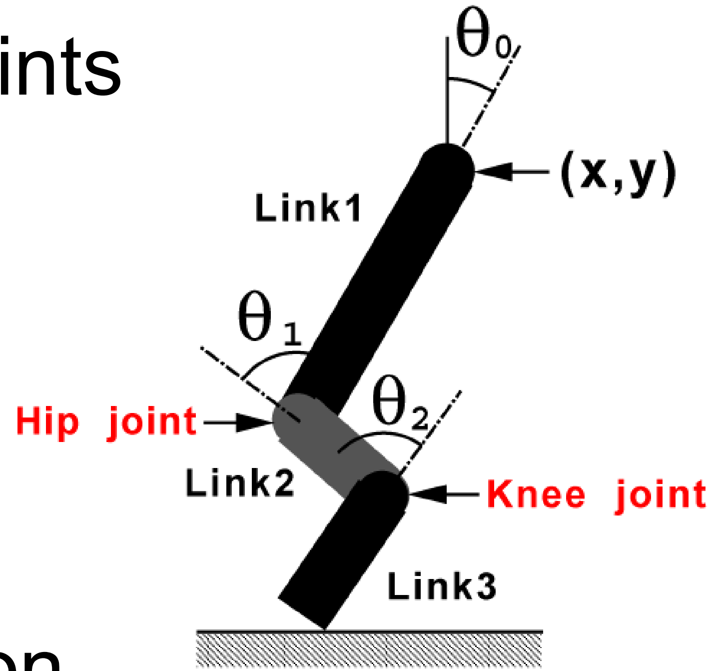
- Similar to supervised learning, the goal is to let the student learn the supervisor's knowledge.
  - However, the supervisor does not answer the student's questions directly.
- Instead, the supervisor gives a reward depending on the student's predicted answer.
  - The student learns to maximize the rewards.
- Similar to children's learning!




# Example: Robot Stand-up Motion



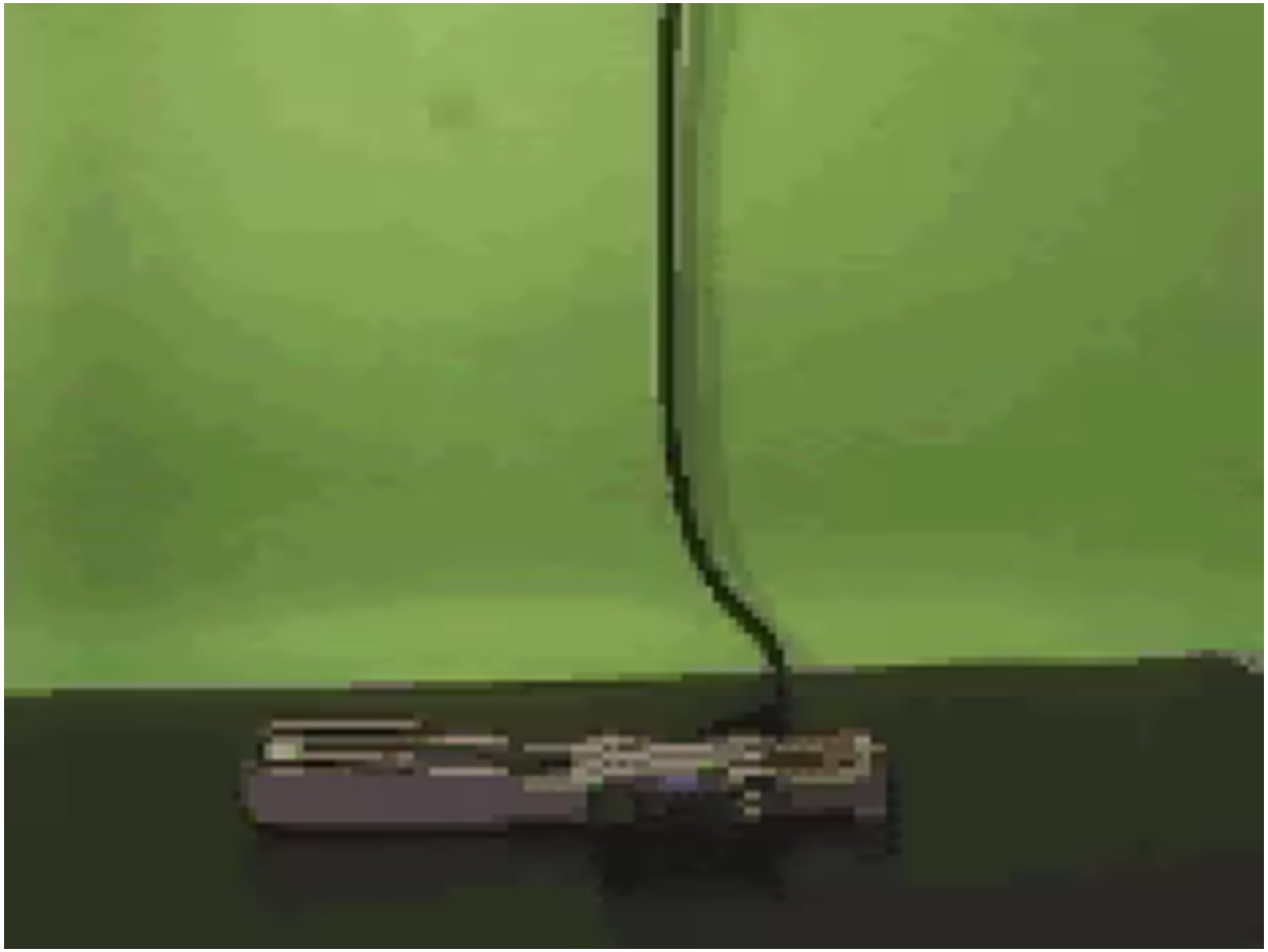
- Toy humanoid robot:
  - Equipped with two controllable joints
- Goal: Learn standing-up motion autonomously
- Reward: Height of the head
  - Robot stands up
    = Its head is at the highest position



# Demonstration



- Before learning



http://www.kawato.jst.go.jp/xmorimo/stup_movie.html

# Demonstration

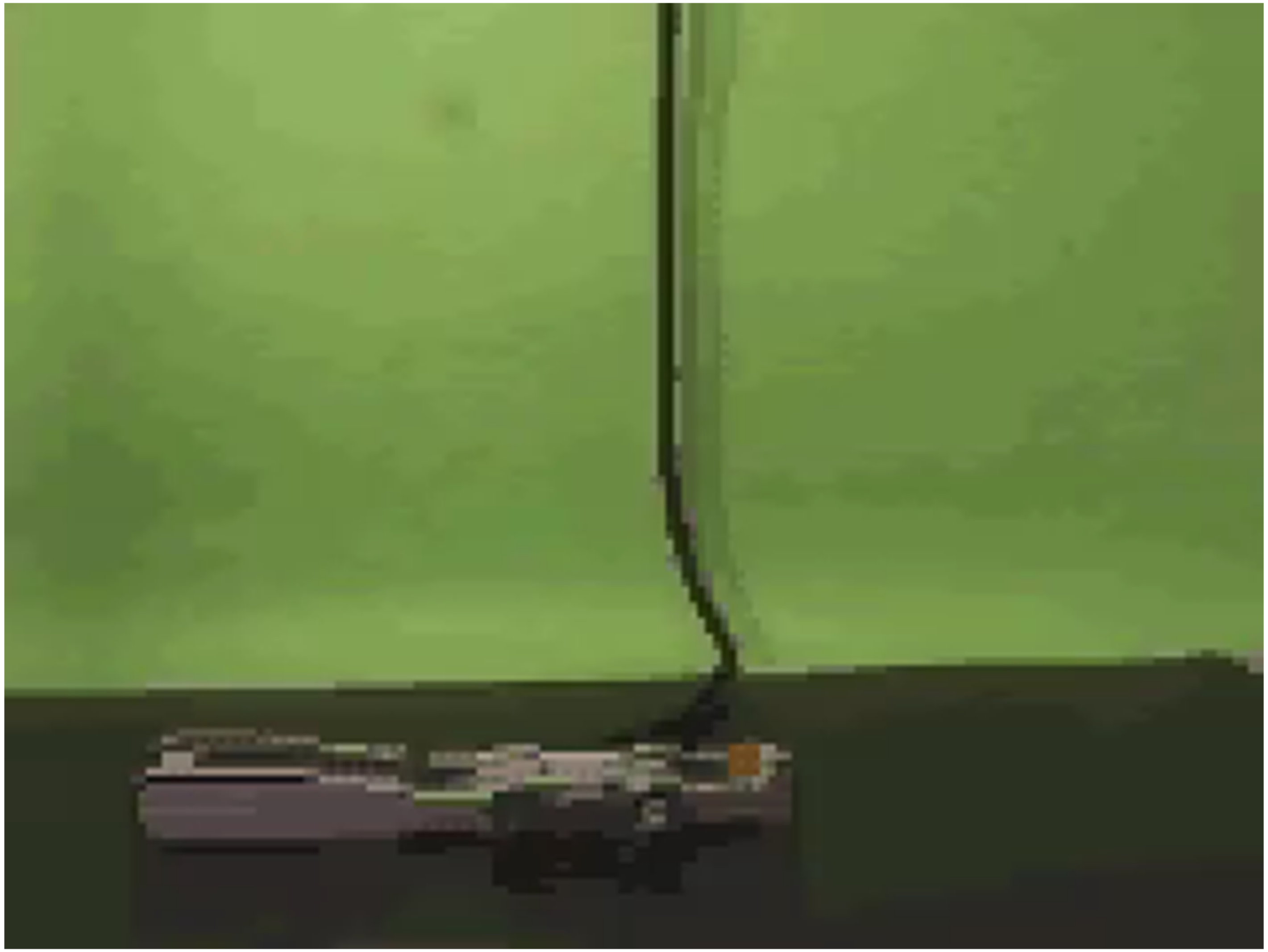
- After 750th training

http://www.kawato.jst.go.jp/xmorimo/stup_movie.html

# Demonstration



- After 920th training



http://www.kawato.jst.go.jp/xmorimo/stup_movie.html

# More Examples


Game AI

https://www.itmedia.co.jp/news/articles/1705/23/news122.html

Robot control

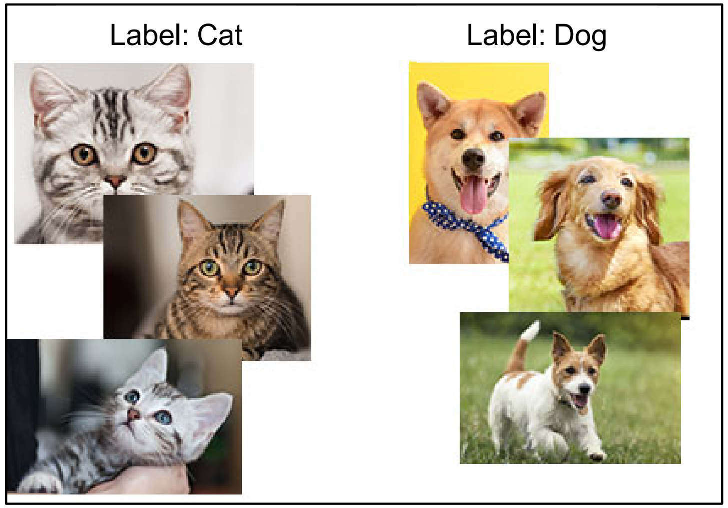
Sugimoto et al. IEEE-RAM2016

# Contents

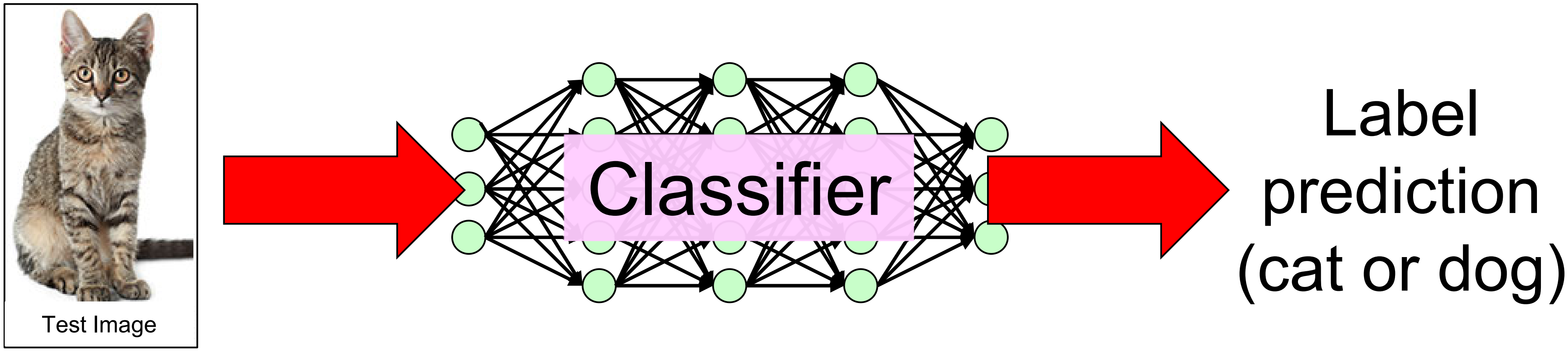
# How Can We Generalize in Supervised Learning?

- Ex: Image classification
  - Based on supervised learning, we develop a classifier that recognizes dogs and cats.

Training images



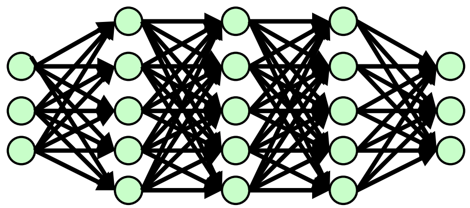

https://www.cct-inc.co.jp/media/ai/news/primer/ai-08/



# Training and Prediction



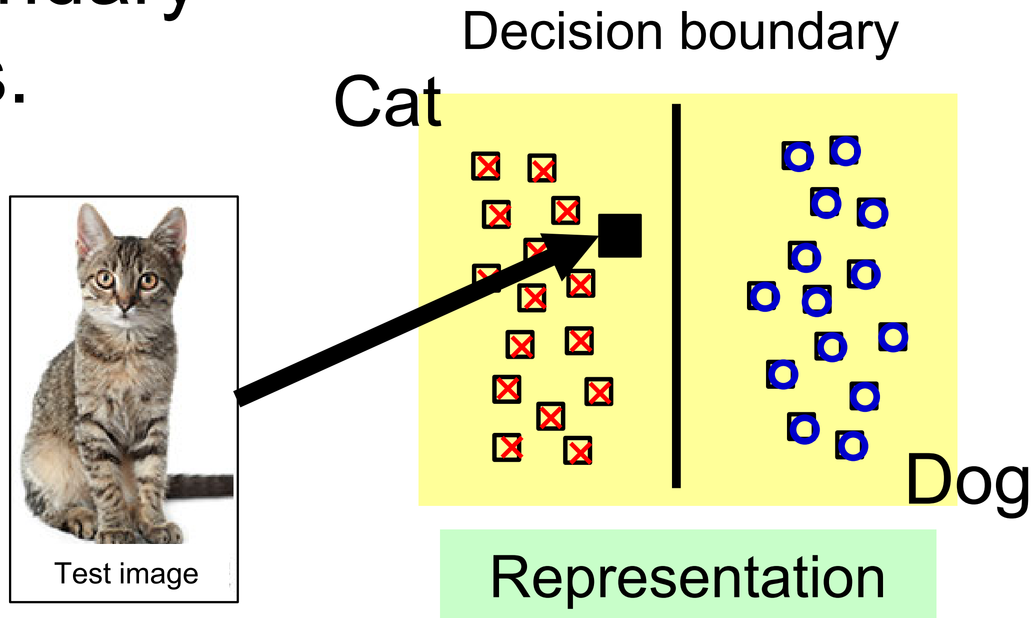
- In the classifier, an image is represented as a point.



- Training:
  - Prepare (many) labeled training images.
  - Learn the decision boundary between cats and dogs.


https://www.cct-inc.co.jp/media/ai/news/primer/ai-08/

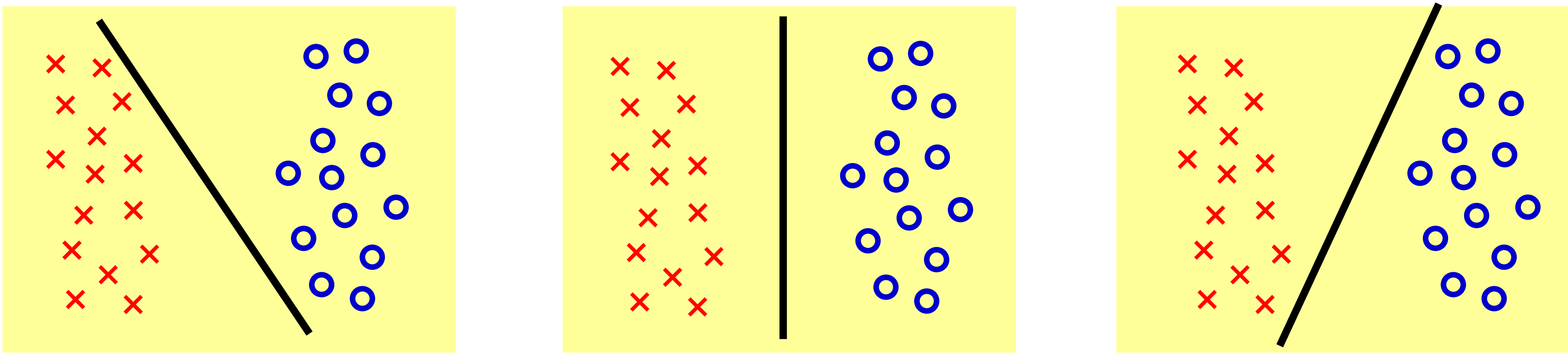
- Prediction:
  - Classify unlearned test images.



# Better Prediction



- It is easy to separate given training data:.

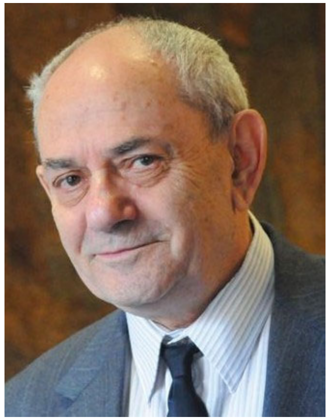


- Which one generalizes the best for unseen test data?
- Since we don't know where test data appear, this question may not be answered naively.
- Use statistical learning theory!

# Statistical Learning Theory

Vapnik & Chervonenkis (1971)

- Prediction error for test data cannot be accessed, but its probabilistic upper bound can!



https://mlatcl.github.io/deepnn/lectures/01-01-introduction.html

With probability $1-\delta$,

$$\text{Prediction Error} \leq \sqrt{\frac{1}{n}\left(V\left(\log\frac{2n}{V}+1\right)+\log\frac{4}{\delta}\right)}$$

$n$ : Number of training data

$V$ : VC-dimension (complexity of a classifier)

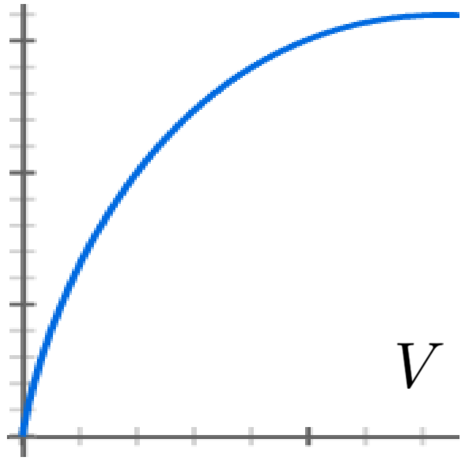
# Statistical Learning Theory



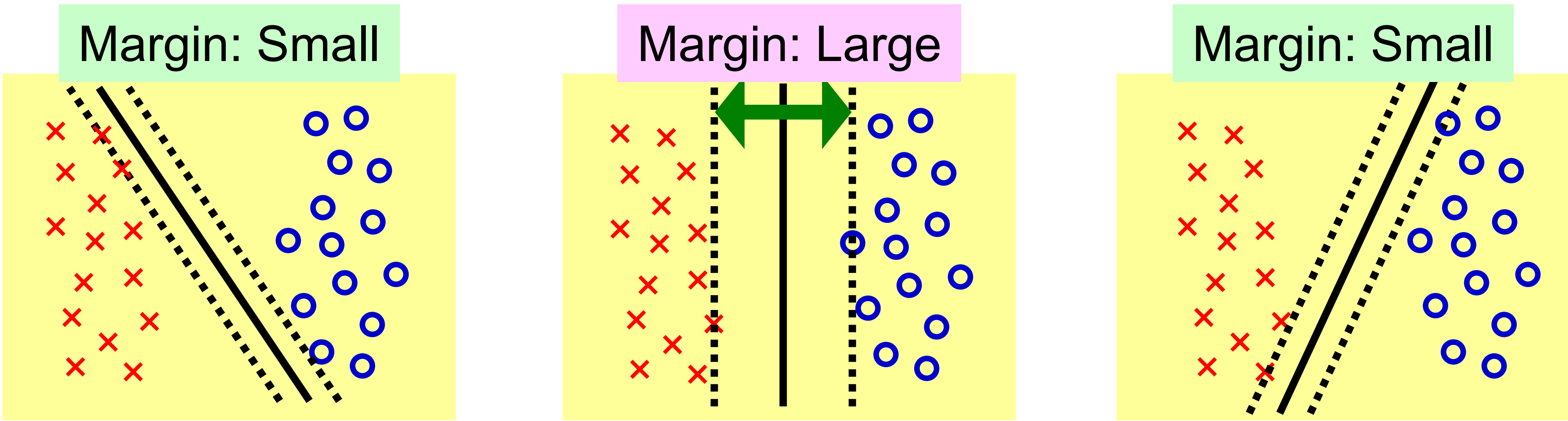
With probability $1-\delta$,

$$\text{Prediction Error} \leq \sqrt{\frac{1}{n}\left(V\left(\log\frac{2n}{V}+1\right)+\log\frac{4}{\delta}\right)}$$

$V$ : VC-dimension (complexity of a classifier)




- Smaller $V$ yields a smaller upper bound.
- A larger margin yields a smaller upper bound of $V$.




- This method is called the support vector machine.

# Contents

## Statistical learning theory for linear classifiers:

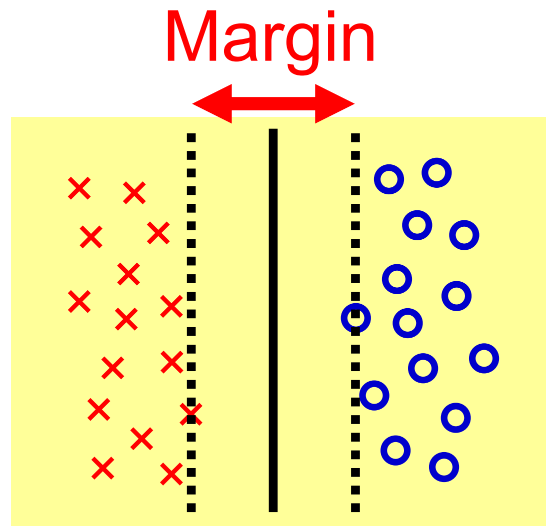
- The maximum-margin solution is the best.

## Further investigation:

A) What if training data are not linearly separable?

B) What if training data contain label noise?

C) What if the number of training data is small?

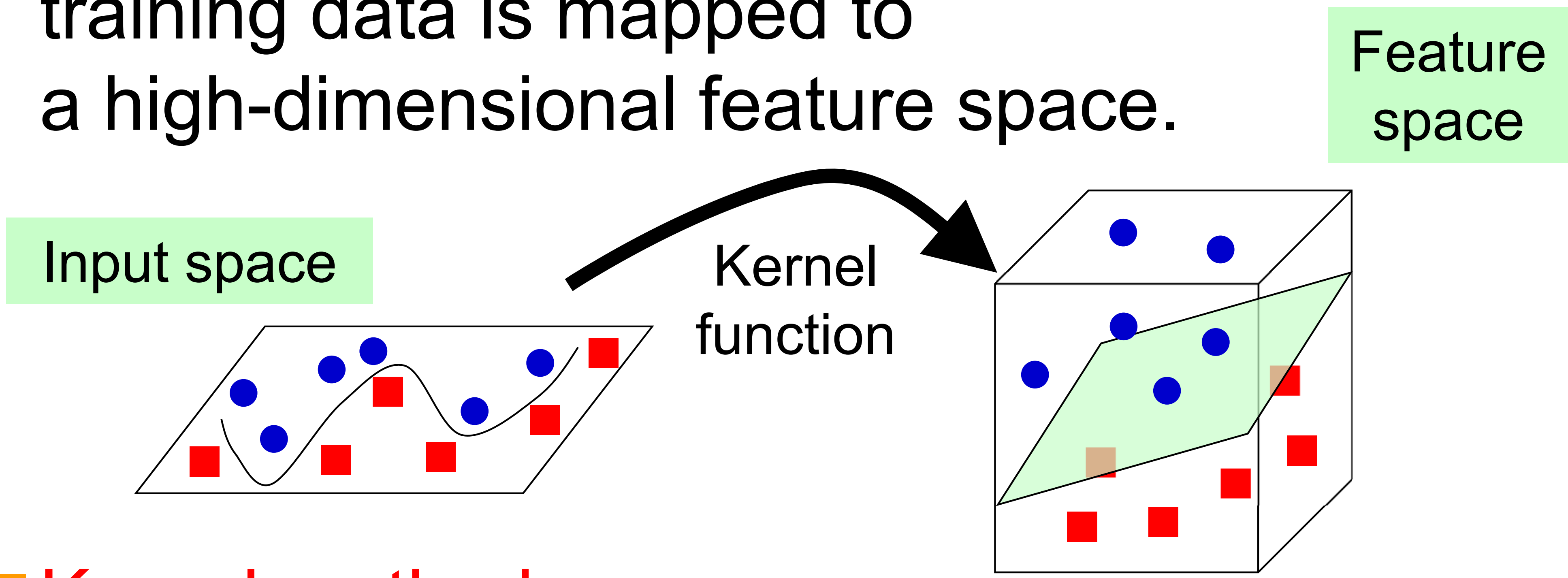



Margin: Large

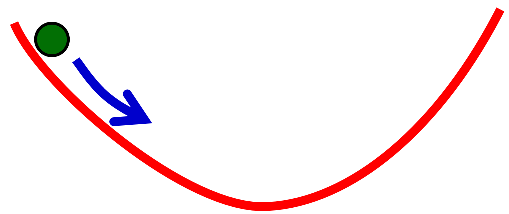
# Non-linear Extension: Kernel Models



- By a pre-fixed non-linear transformation, training data is mapped to a high-dimensional feature space.




- Kernel methods:

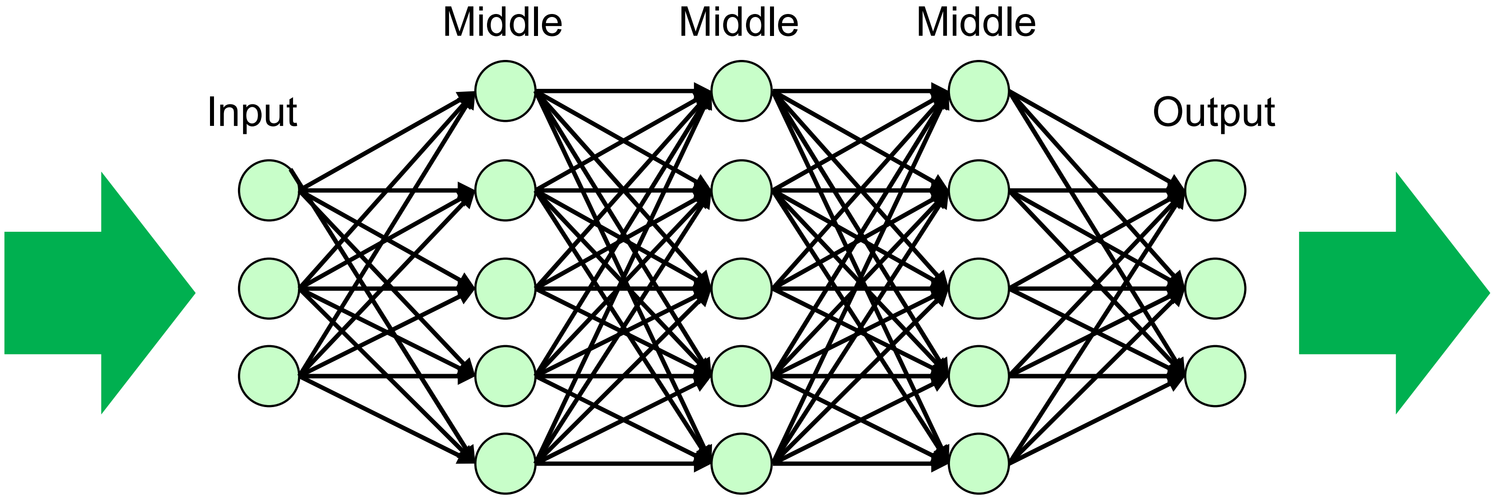


  - Model training is simple (convex)
  - Work moderately well for high-dimensional data.
  - Slow for large-scale data.

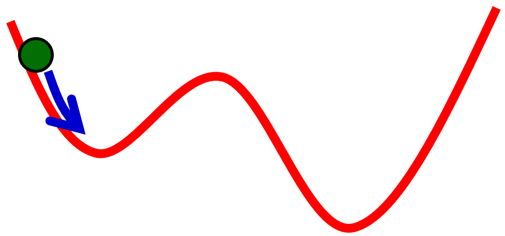
# Non-linear Extension: Deep Models



- Repeat non-linear transformations many times to represent complex functions



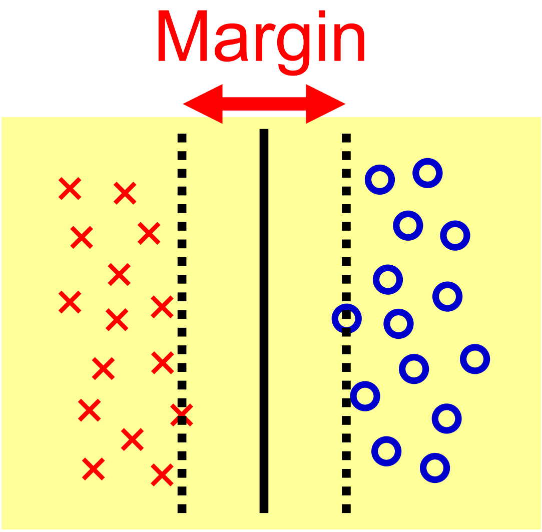

- Deep learning:
  - Model training is complex (non-convex), but is significantly improved.
  - Works well for high-dimensional data.
  - Efficient trainable for large-scale data.



# Advances in Supervised Classification

- Statistical learning theory for linear classifiers:
  - The maximum-margin solution is the best.

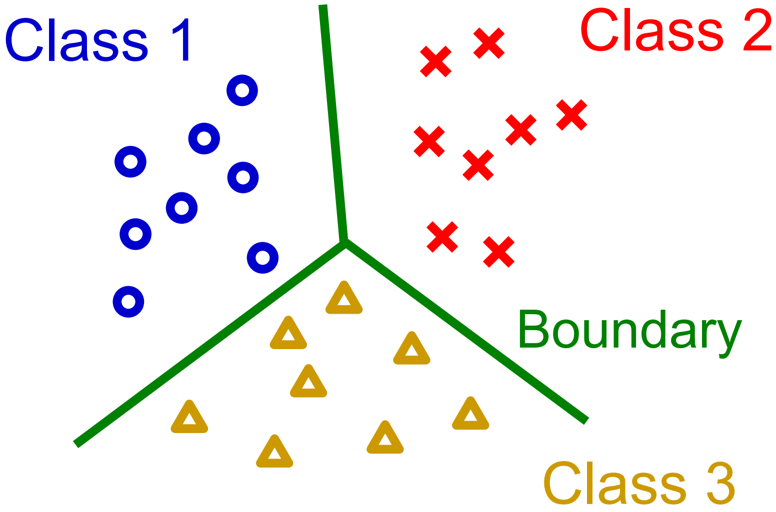
- Further investigation:
  - A) What if training data are not linearly separable?
  - B) What if training data contain label noise?
  - C) What if the number of training data is small?




Margin: Large

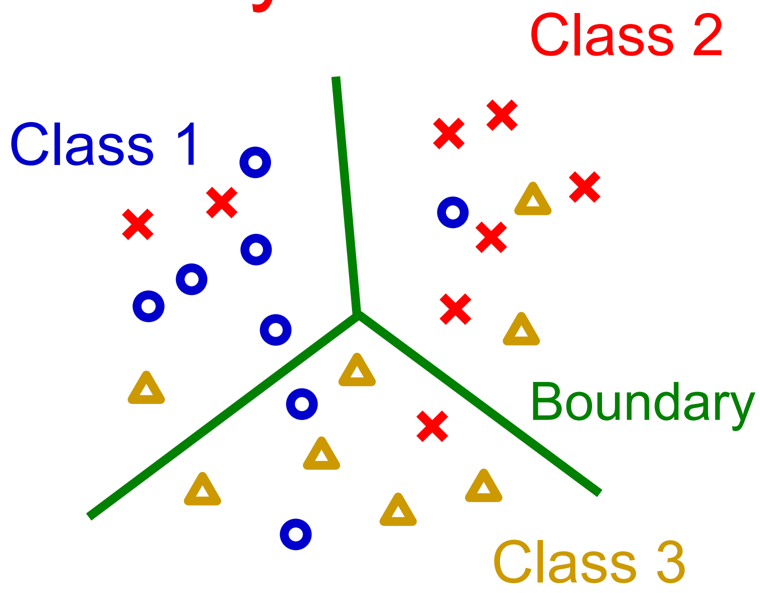
# Noisy Labels

- Supervised classification with clean labels:




Naïve large-margin solution works well.

- However, real-world labels are noisy possibly due to human error:




Large-margin solution suffers from severe overfitting.

- the influence of noise remains even with infinite training data

# Noisy Labels



- **Noise transition matrix $T$:**

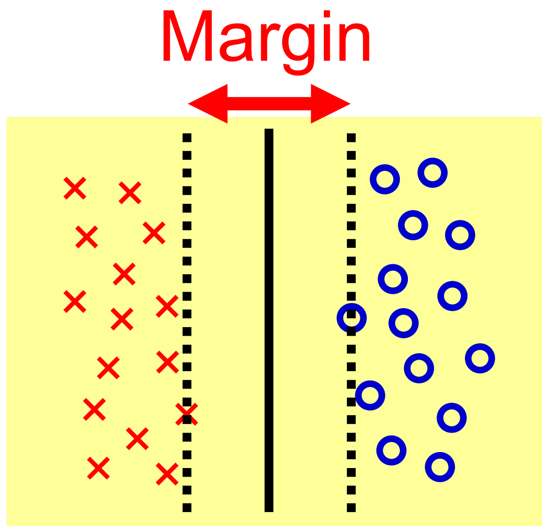
$$T^\top = \begin{array}{|c|c|c|}\hline 1 & 0.1 & 0.5 \\ \hline 0 & 0.8 & 0.5 \\ \hline 0 & 0.1 & 0 \\ \hline\end{array}$$

  - Clean-to-noisy flipping probability.
- Major approaches: Patrini et al. (CVPR2017)
  - Loss correction by $T^{-1}$ to eliminate noise.
  - Classifier adjustment by $T^\top$ to simulate noise.
- We want to estimate $T$ only from noisy data:
  - Use human cognition as a "mask" for $T$. Han et al. (NeurIPS2018)
  - Learn $T$ and a classifier dynamically. Xia et al. (NeurIPS2019)
  - Decompose $T$ into simpler components. Yao et al. (NeurIPS2020)
  - Regularize $T$ to be estimable. Zhang et al. (ICML2021), Li et al. (ICML2021)
  - Extension to input-dependent noise $T(x)$.

Xia et al. (NeurIPS2020), Berthon et al. (ICML2021), Cheng et al. (CVPR2022)

- Statistical learning theory for linear classifiers:
  - The maximum-margin solution is the best.

- Further investigation:
  - A) What if training data are not linearly separable?
  - B) What if training data contain label noise?
  - C) What if the number of training data is small?

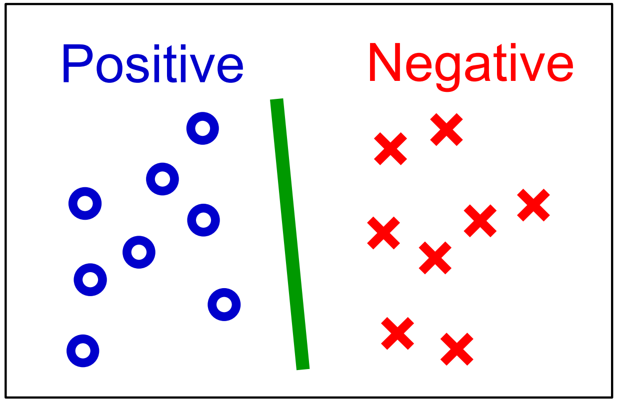



Margin: Large

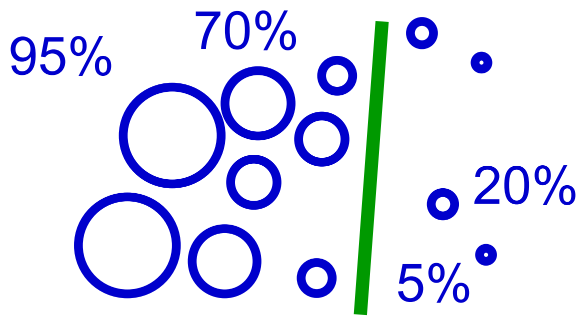
# Weakly Supervised Classification



- Use “weak” labels that can be collected easily:

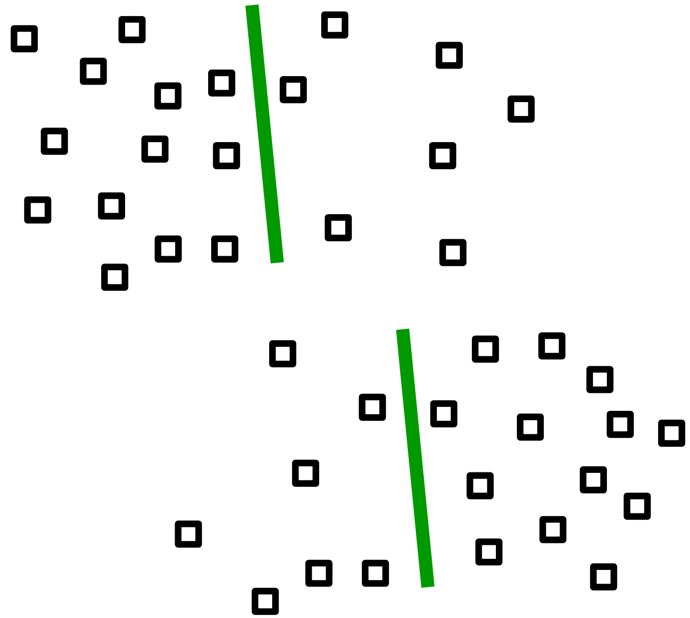



Positive-confidence



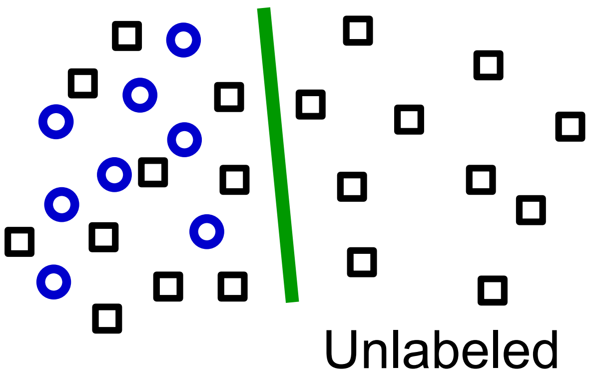

Purchase prediction

Unlabeled-Unlabeled



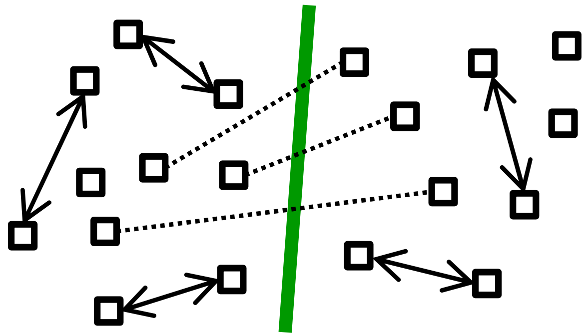
Different populations

Positive-Unlabeled



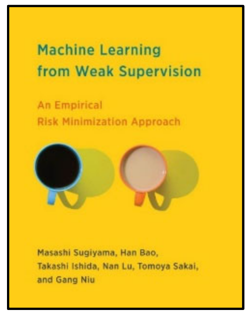

Click prediction

Similar-Dissimilar



Sensitive prediction

Sugiyama, Bao, Ishida, Lu, Sakai & Niu, Machine Learning from Weak Supervision. MIT Press, 2022.



# Contents

# Lecture Schedule (2025S)



1. ~~04/8 Introduction~~
2. 04/15 Regression 1
3. 04/22 Regression 2
- 04/30 Cancelled
4. 05/13 Classification 1
5. 05/20 Classification 2
6. 05/27 Deep learning 1
- 06/03 No lecture
7. 06/10 Deep learning 2
8. 06/17 Deep learning 3
9. 06/24 Semi-supervised learning
10. 07/01 Language models
11. 07/08 Representation learning 1
12. 07/15 Representation learning 2
13. 07/22 Advanced topics

- Zoom URL: https://u-tokyo-ac-jp.zoom.us/j/82403666641?pwd=KRS6iPnUXrLn3Aa79Qe3fdybxADwkS.1
- UTOL: https://utol.ecc.u-tokyo.ac.jp/lms/course?idnumber=2025_4740_47130-53_01

# Homework 1

- The probability of Liking/Hating this course is

$$p(X = \mathrm{L}) = 0.8 \qquad p(X = \mathrm{H}) = 0.2$$

- Among those who like/hate this course, the probabilities of feeling Sleepy/Awake during the class are

$$p(Y = \mathrm{S}|X = \mathrm{L}) = 0.25$$

$$p(Y = \mathrm{S}|X = \mathrm{H}) = 0.25$$

- Answer the following questions:
  A) Find $p(X = \mathrm{L}, Y = \mathrm{S})$
  B) Find $p(Y = \mathrm{S})$
  C) Find $p(X = \mathrm{L}|Y = \mathrm{S})$
  D) Are “liking/hating this course” and “feeling sleepy/awake during the class” statistically independent?

# Homework 2



- Prove the following equations:

A) $\mathbb{E}(c) = c$

B) $\mathbb{E}(X + c) = \mathbb{E}(X) + c$

C) $\mathbb{E}(cX) = c\mathbb{E}(X)$

$\mathbb{E}$ : Expectation
$c$ : Constant
$X$ : Random variable

# Homework 3

- Prove the following equations:
  - A) $\mathbb{V}(c) = 0$
  - B) $\mathbb{V}(cX) = c^2\mathbb{V}(X)$
  - C) $\mathbb{V}(X + c) = \mathbb{V}(X)$

$\mathbb{V}$ : Variance
$c$ : Constant
$X$ : Random variable

# Homework 4



- Prove the following equations:

A) $\mathbb{E}(X + X') = \mathbb{E}(X) + \mathbb{E}(X')$

B) $\mathbb{V}(X + X') = \mathbb{V}(X) + \mathbb{V}(X') + 2\mathbb{C}(X, X')$

$X, X'$ : Random variables
$\mathbb{C}$ : Covariance

$$\mathbb{C}(X, X') = \mathbb{E}[(X - \mathbb{E}[X])(X' - \mathbb{E}[X'])]$$

# Submission of Homework



- Upload your homework to UTOL before the beginning of the next class.

https://utol.ecc.u-tokyo.ac.jp/lms/course?idnumber=2025_4740_47130-53_01

- Late submissions will be accepted, but scores will be reduced.

# Caution

- Plagiarism is strictly prohibited:
  - All the grades may be invalidated if plagiarism is found in one of your reports.
- On the use of AI tools such as ChatGPT:
  - Do not use them for text generation from scratch.
  - You may use them to improve your own writing (e.g., for correcting grammatical errors), but must declare how you used them in your report.
  - In any case, you have to take responsibilities on your own report; note that the text generated by AI tools may contain plagiarism.

# 注意

- 剽窃行為は固く禁じられています:
  - レポートに盗用が発見された場合，全ての成績が無効になることがあります．
- ChatGPTなどのAIツールの使用について
  - ゼロから文章を生成するために使用しないこと．
  - 自分の文章をより良くするため（文法的な誤りを修正するためなど）に使用してもよいですが，どのように使用したかをレポートに記載する必要があります．
  - AIツールで生成された文章には剽窃が含まれている可能性があるので，自分のレポートには自分で責任を持つ必要があります．