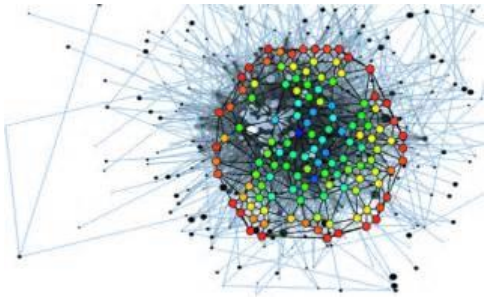
Machine Learning

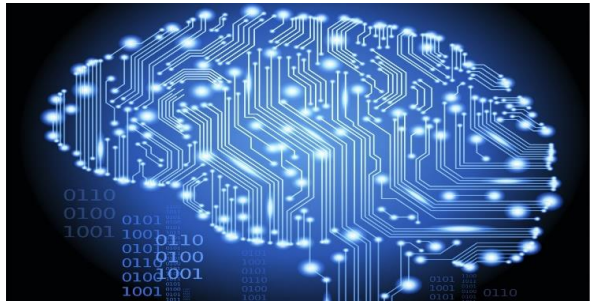
Lecture 1 – Philosophy of Machine Learning

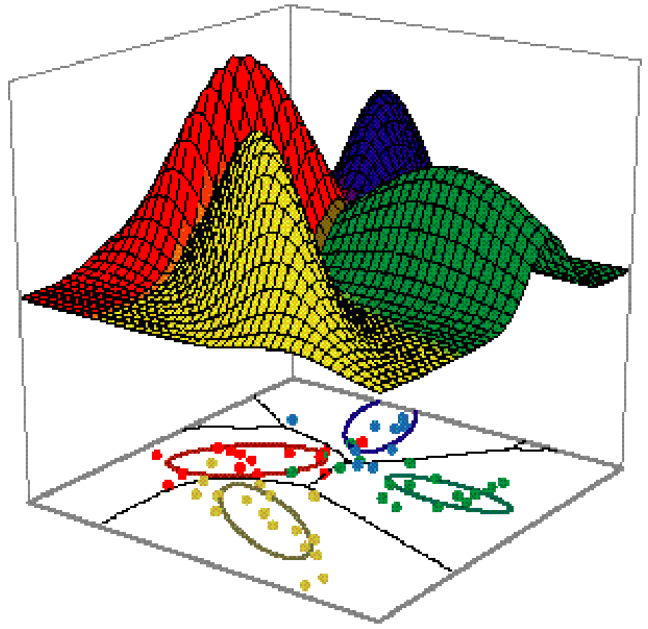
Zengchang Qin (PhD)
Intelligent Computing and Machine Learning Lab
Beihang University, Beijing, China
zengchang.qin@gmail.com

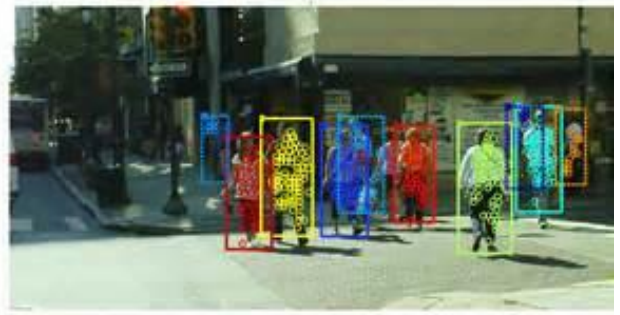
Artificial Intelligence and Machine Learning

- Artificial Intelligence
- Pattern Recognition
- Machine Learning
- Data Mining (Big Data)
- Statistical Learning
- Deep Learning
- Natural Language Processing
- Information Retrieval
- Computer Vision
- Image Processing
- Speech Recognition
- Biometrics
- Bioinformatics



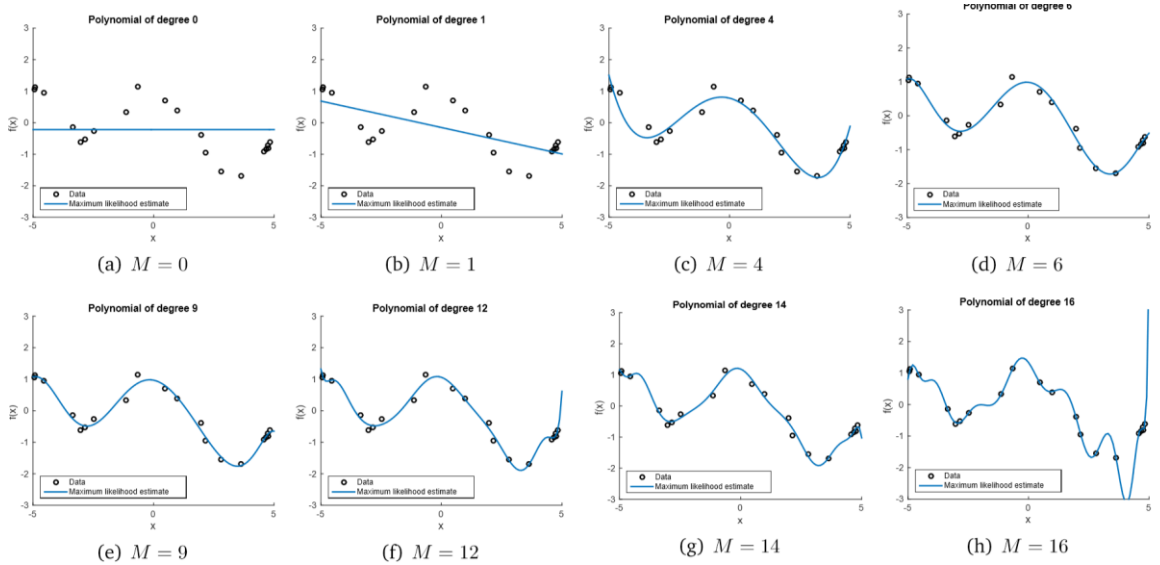






Science of Algorithms

- If machine learning is a science, it is a science of algorithms. - Pat Langley [1]
- Like science, it is about hypothesis testing using evidence.
- The central idea, already proposed by Simon [2] was that the purpose of learning is to improve performance on some class of tasks.
- The UCI Machine Learning Repository is available to the community by FTP in 1987 [1].

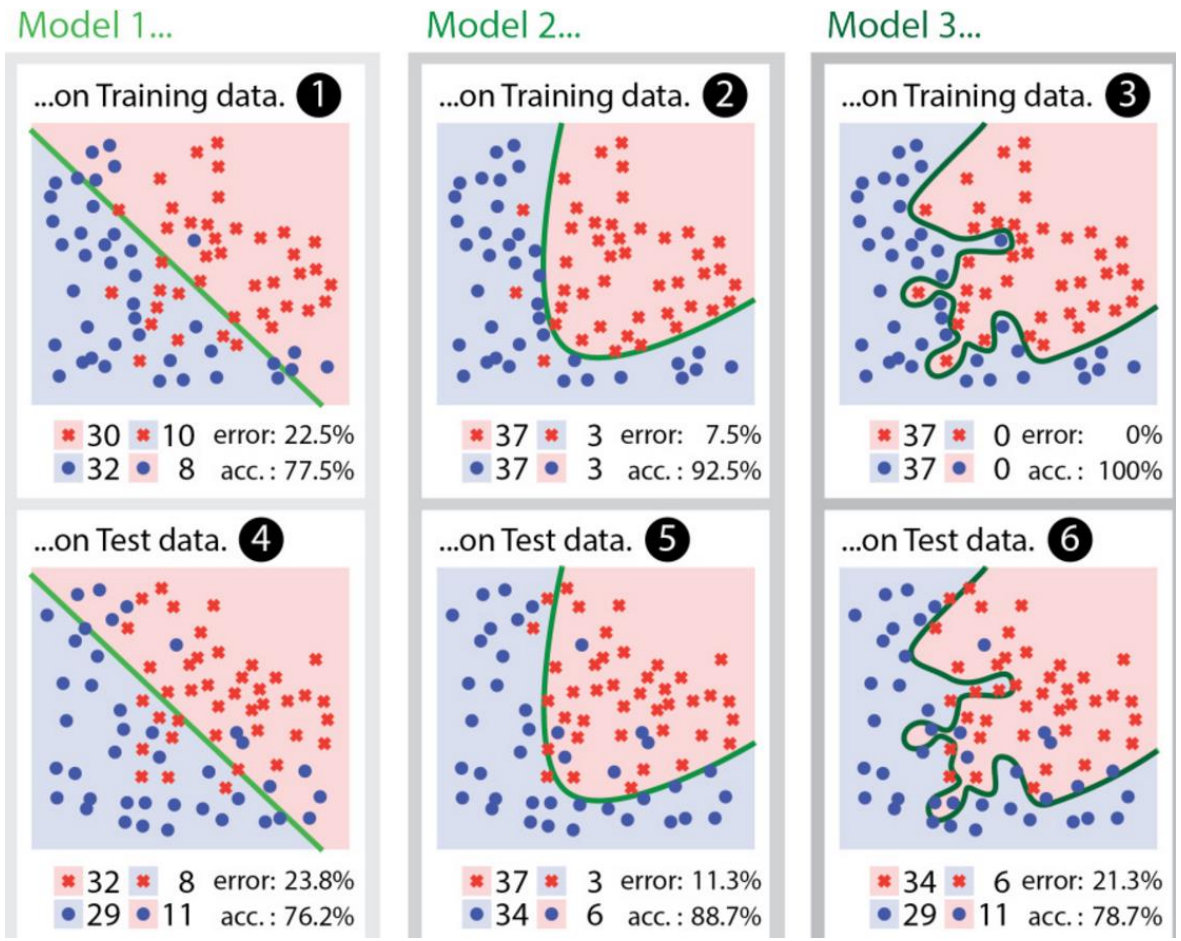


(a) $M = 0$ (b) $M = 1$ (c) $M = 4$ (d) $M = 6$

(e) $M = 9$ (f) $M = 12$ (g) $M = 14$ (h) $M = 16$

[1] http://archive.ics.uci.edu/ml/

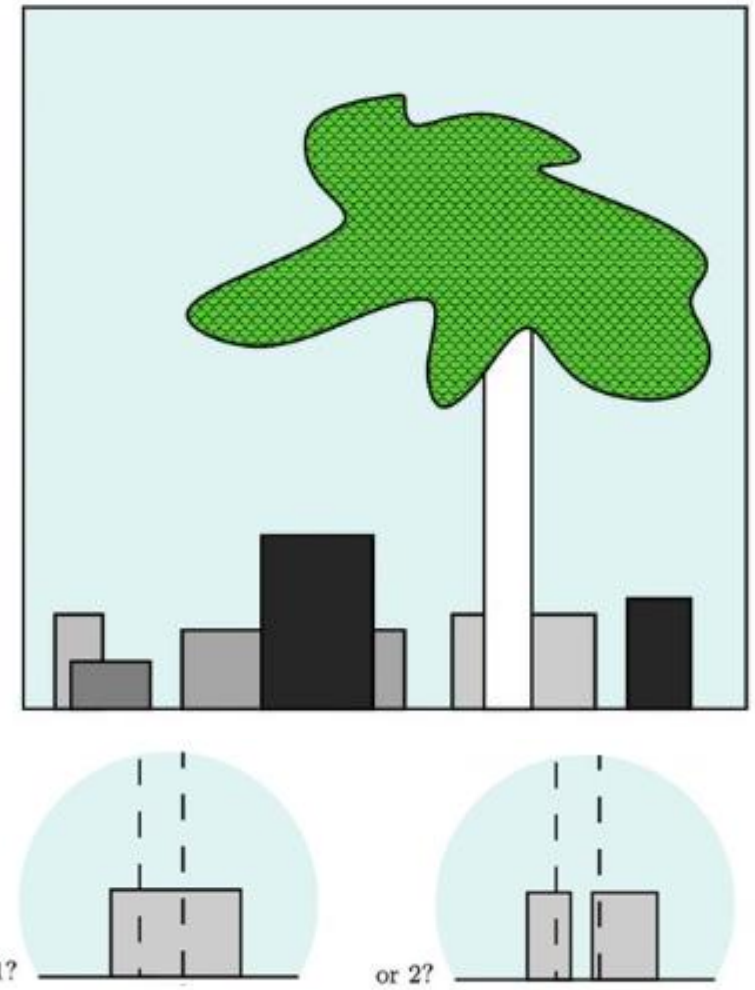
Search in Hypothesis Space

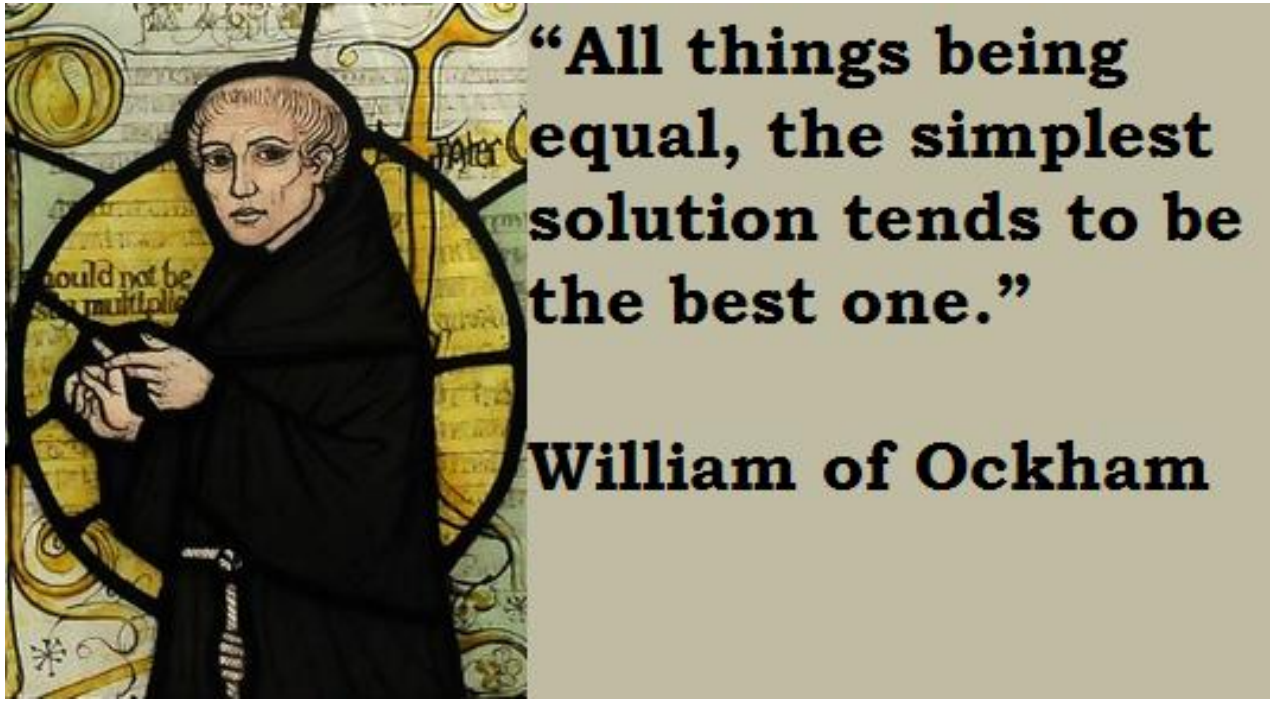


■ A tradeoff between generalization and specification. (E.g, there is no two leave are exactly the same, or all greens are leaves.)

"I've narrowed it to two hypotheses: it grew or we shrunk."

Occam’s Razor



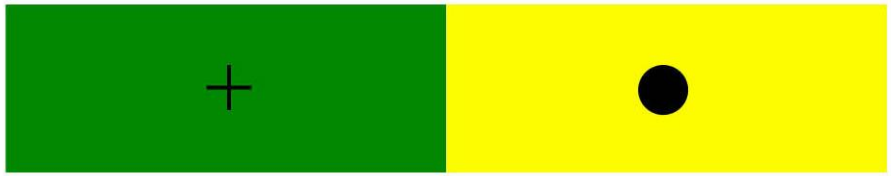


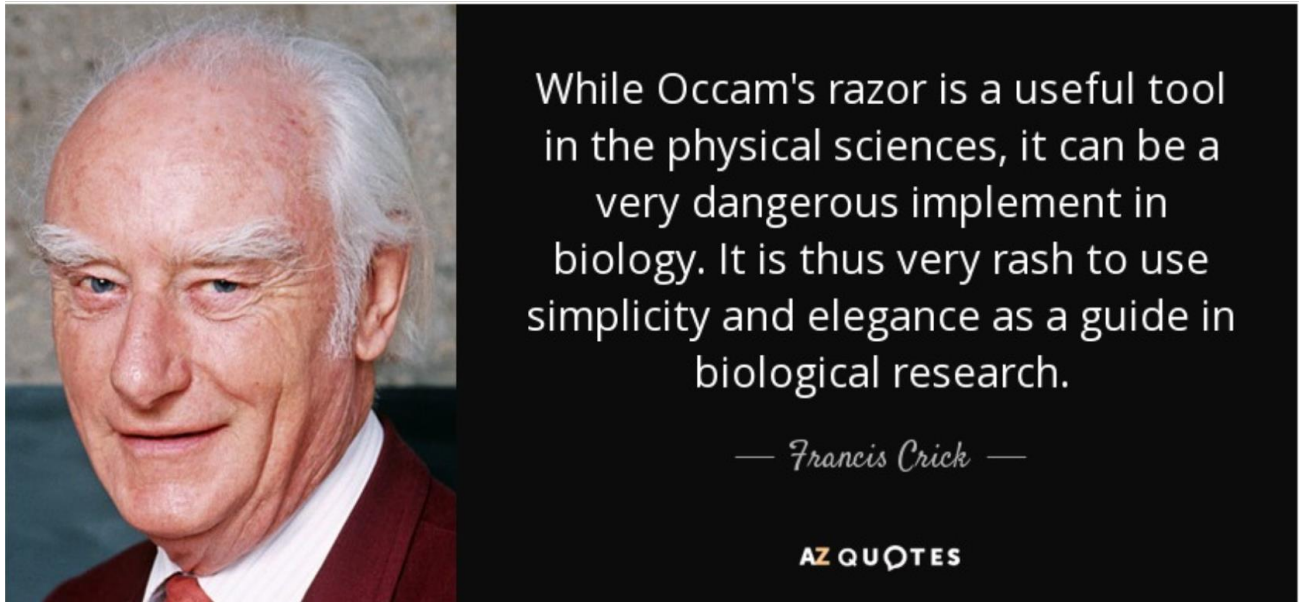
William of Ockham (1287–1347)

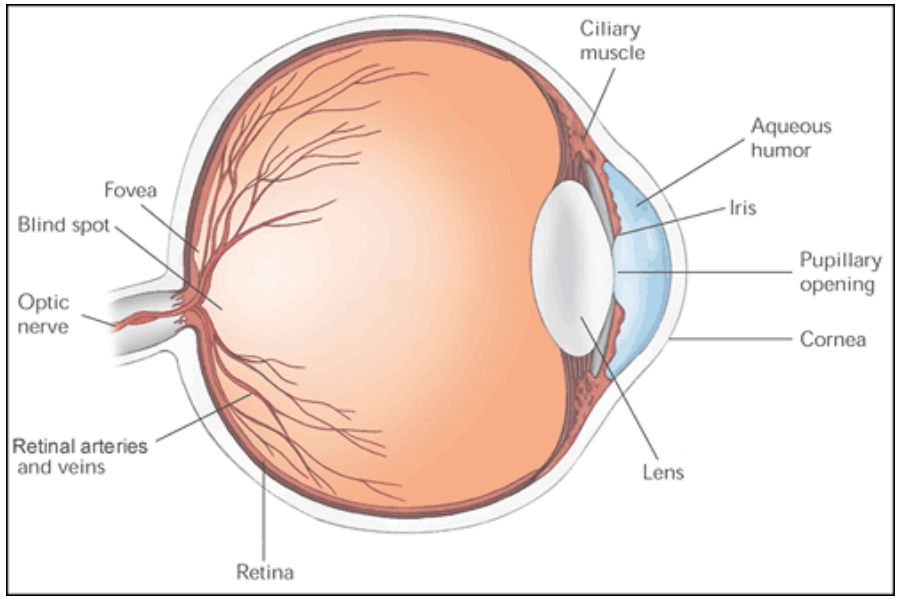
Occam’s Razor also known as the **principle of parsimony** , It is generally understood in the sense that with competing theories or explanations, the simpler one

Danger of Occam's Razor

■Dangerous to biological or social research where innate relations are indeed complex by nature (e.g. Blind point of human eyes.)





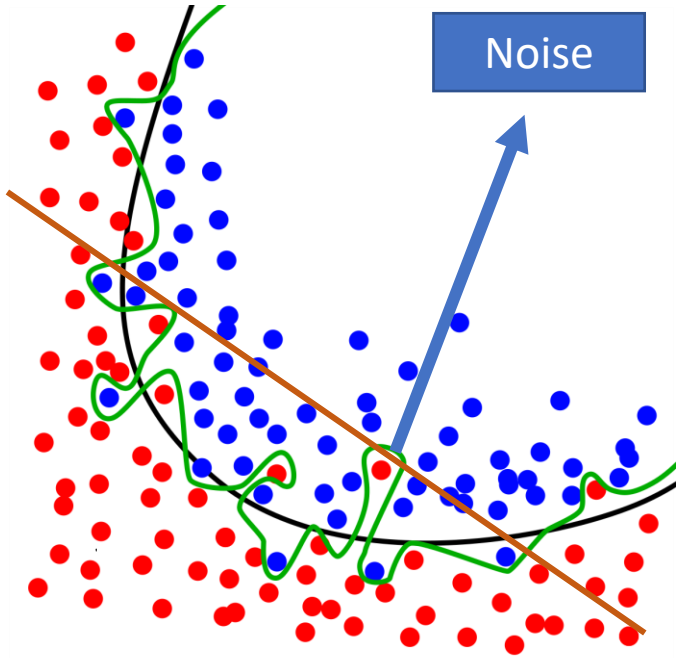


Q: Any other examples of Occam's Razor?

Overfitting

■ Let's say you attend a symphony and want to get the clearest, most faithful sound possible. So you buy a super-sensitive microphone and hearing aid to pick up all the sounds in the auditorium.

■ **Then you start "overfitting," hearing the noise on top of the symphony.** You hear your neighbors shuffling in their seats, the musicians turning their pages, and even the swishing of the conductor's coat jacket.



Q: Any other examples of overfitting?

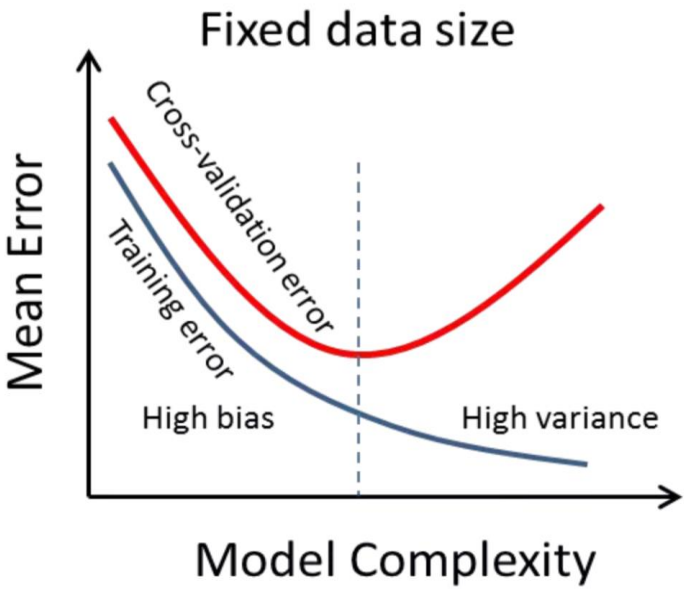
Avoid Overfitting

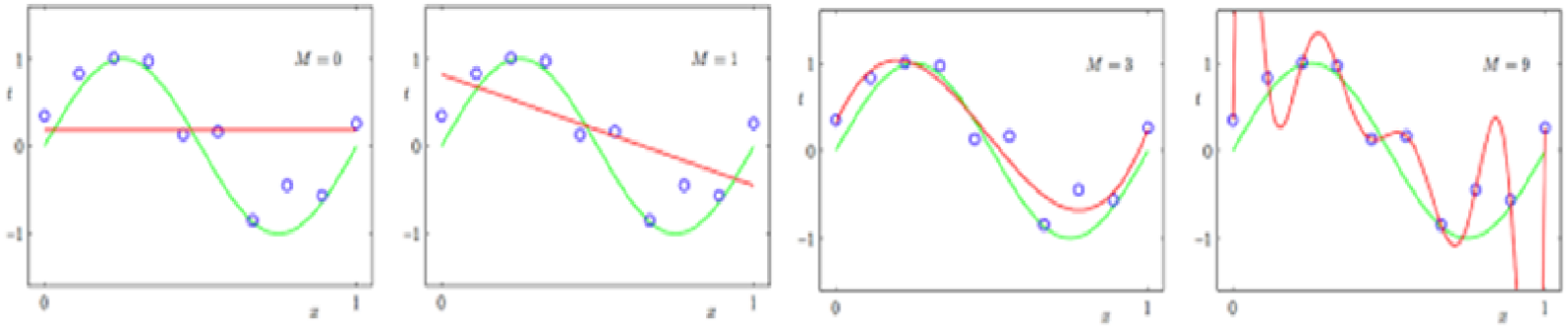
In short, the general strategies are to

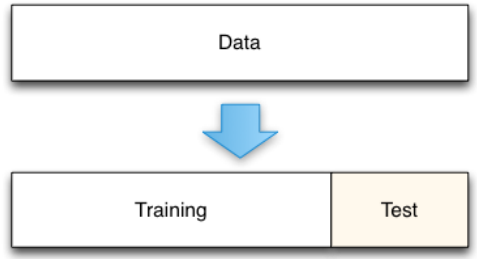
1.Collect more data
2.Use ensemble methods that “average” models
3.Regularization to penalize complexity

$$W = \sum V(f(x_i), t_i) + \lambda \Omega(f)$$

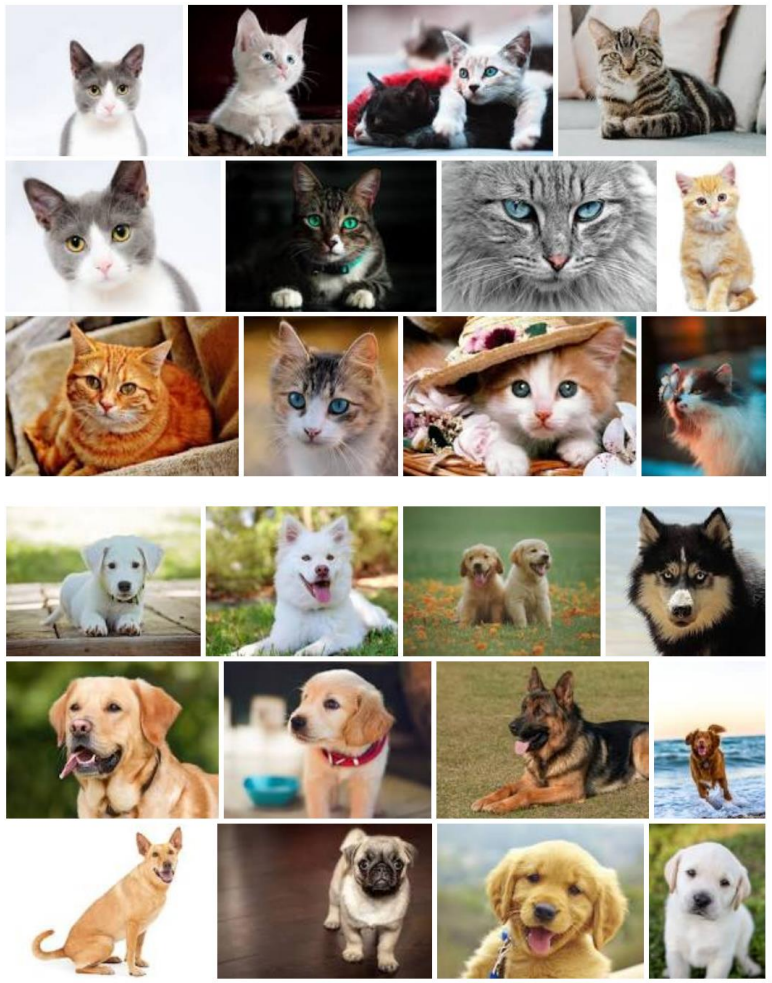
4. N-fold cross-validation

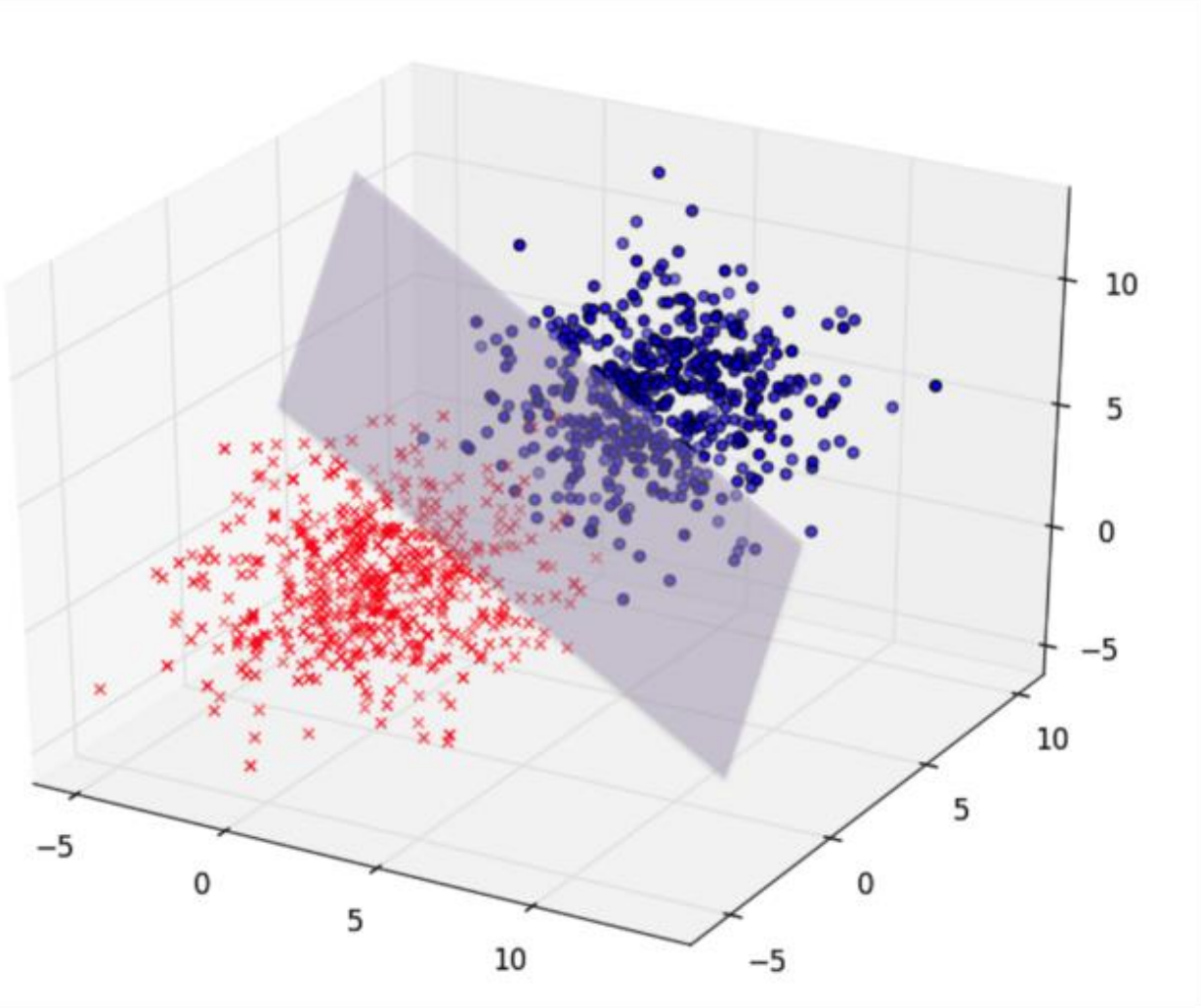






Representation

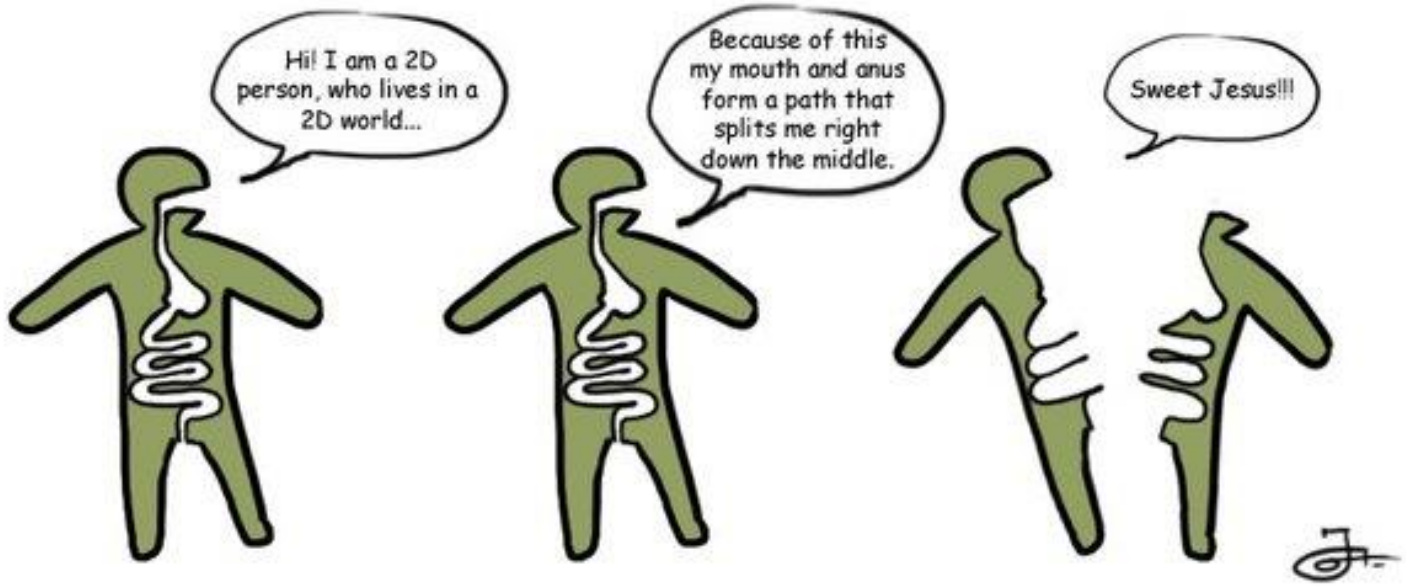


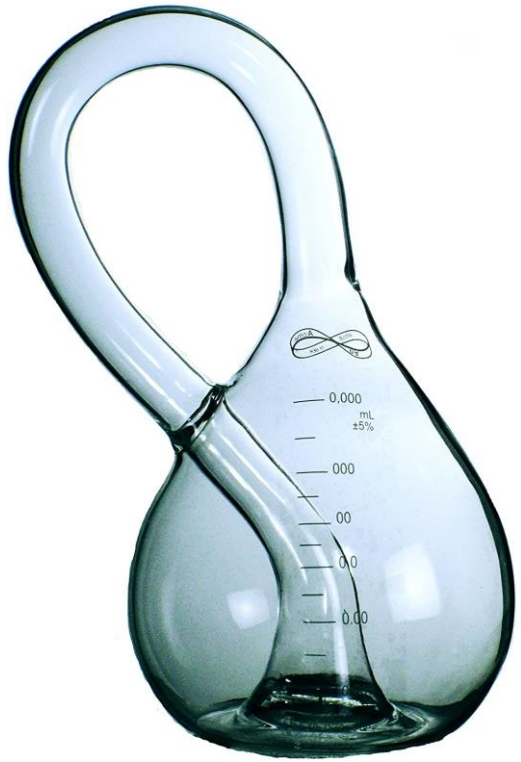


Dimensionality

What being could possess such God-like power?
The answer: a being from a **higher-dimensional** world.

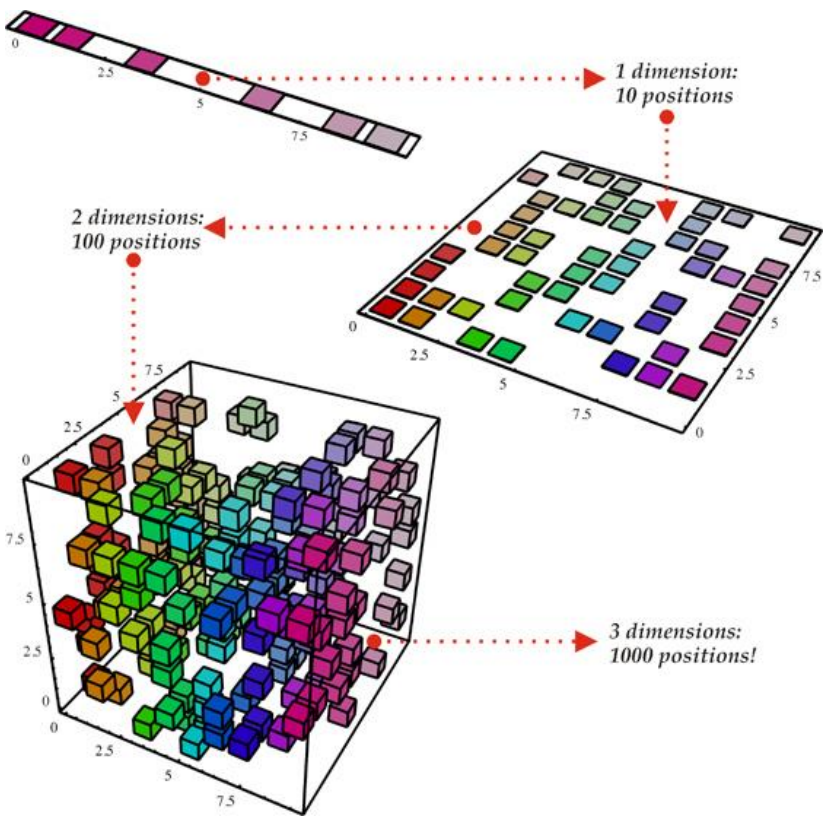
Curved Space

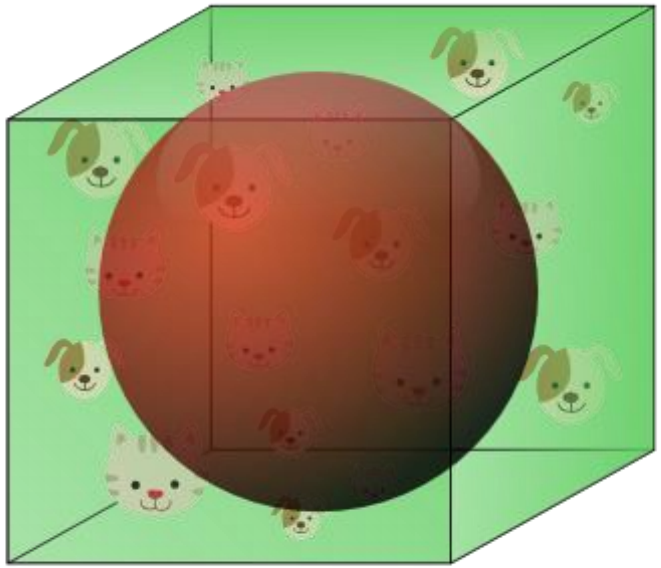


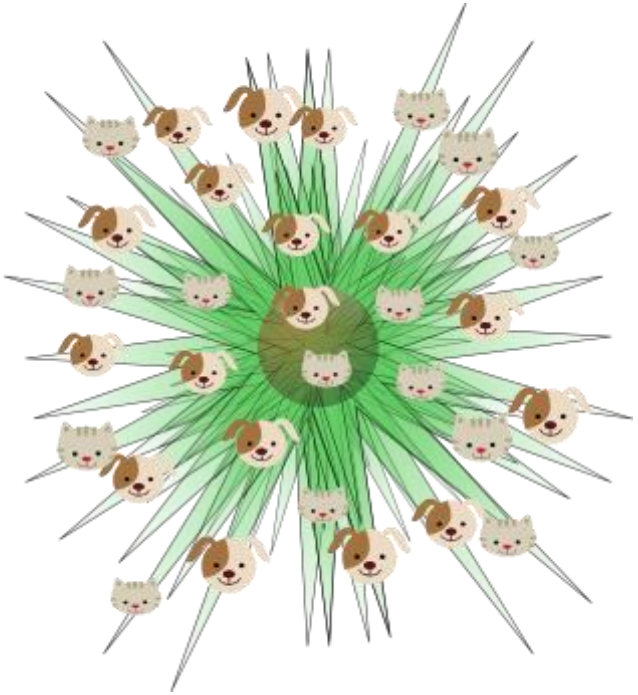


Rudy Rucker, The Fourth Dimension: Toward a Geometry of Higher Reality

Curse of Dimensionality

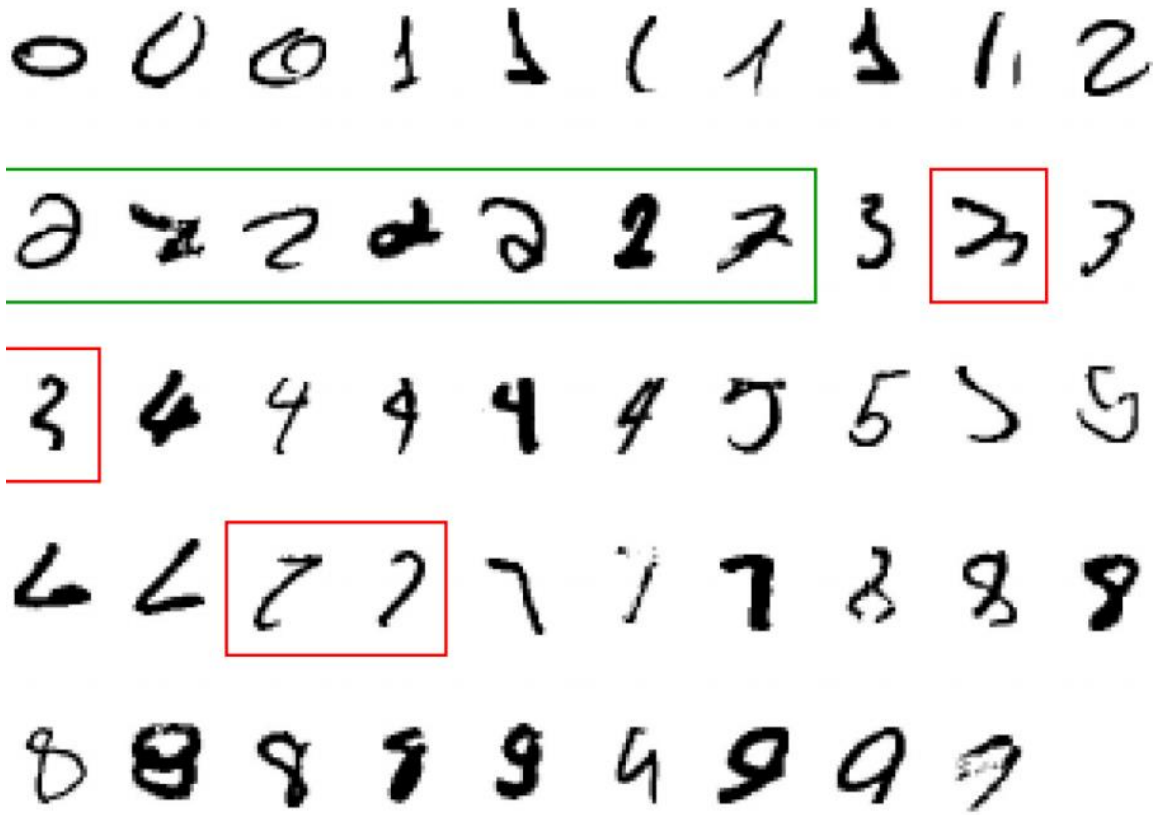




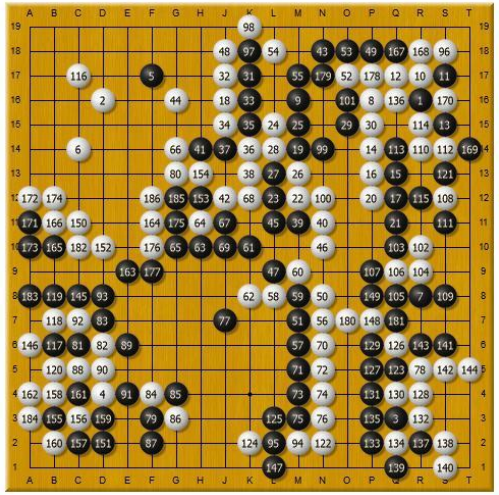


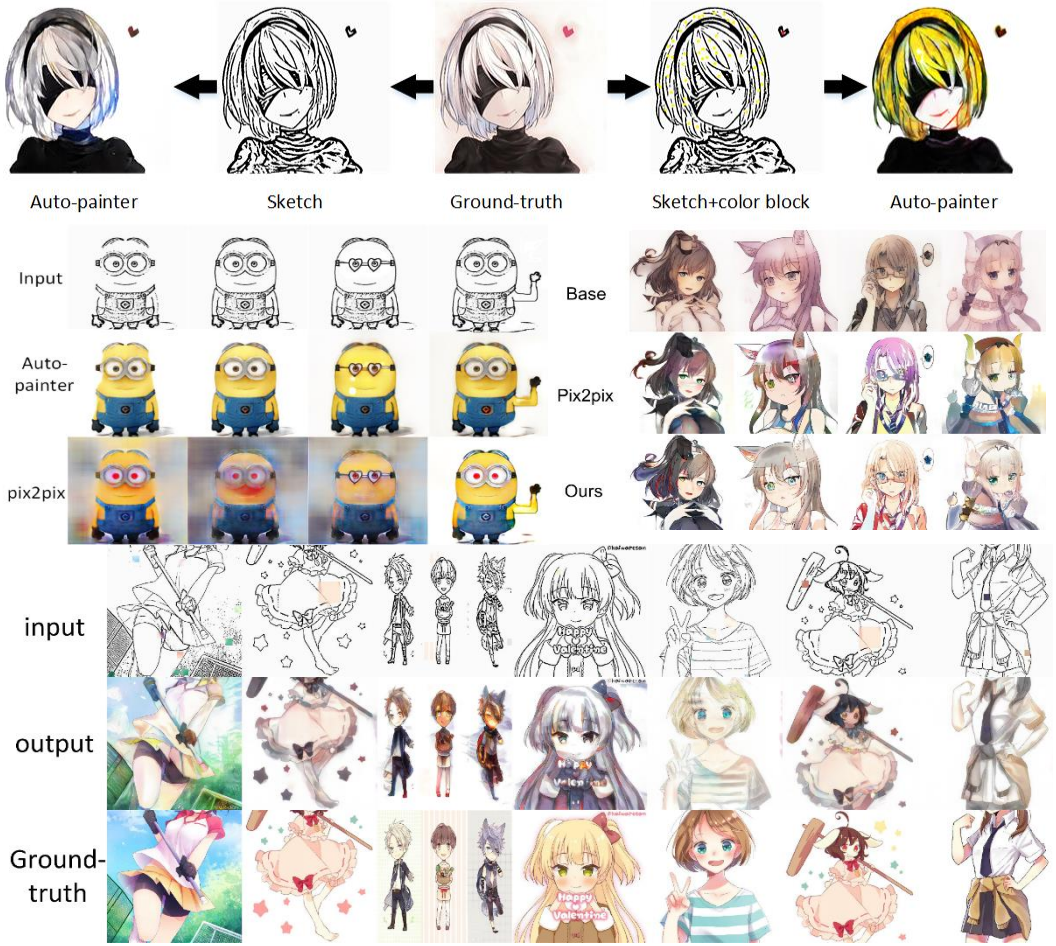
Pattern Recognition

- Instead of writing a program by hand, we collect lots of examples that specify the correct output for a given input.
- A machine learning algorithm then takes these examples and produces a program that does the job.
- We don't have any 'physical' machines to do the job, instead, they are 'mathematical models'



Computer Vision





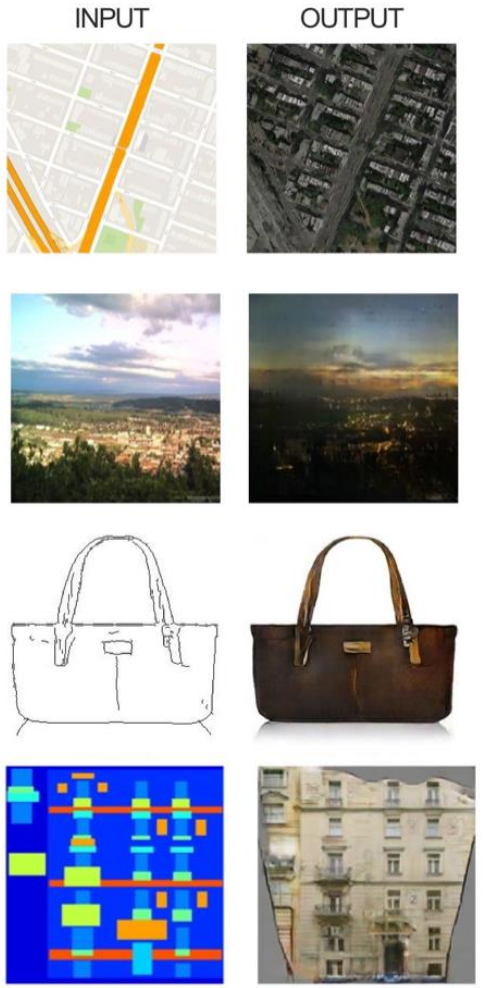
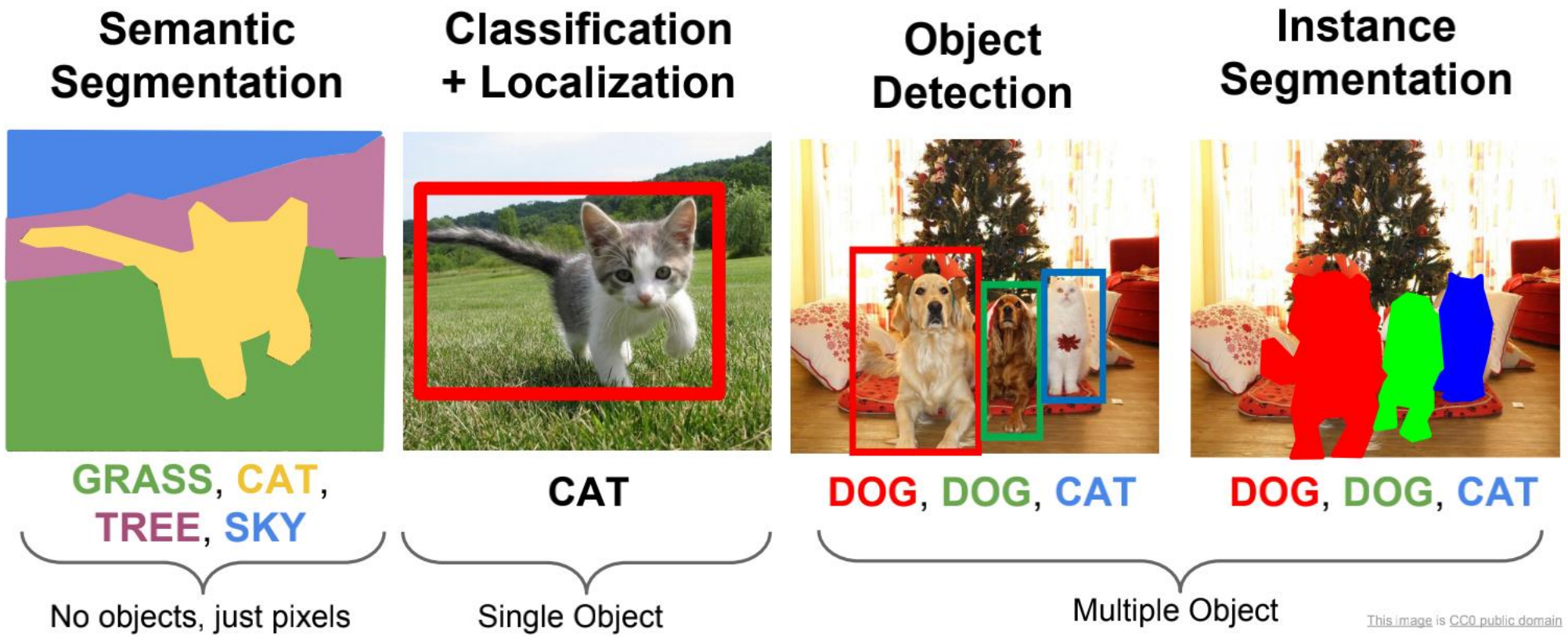
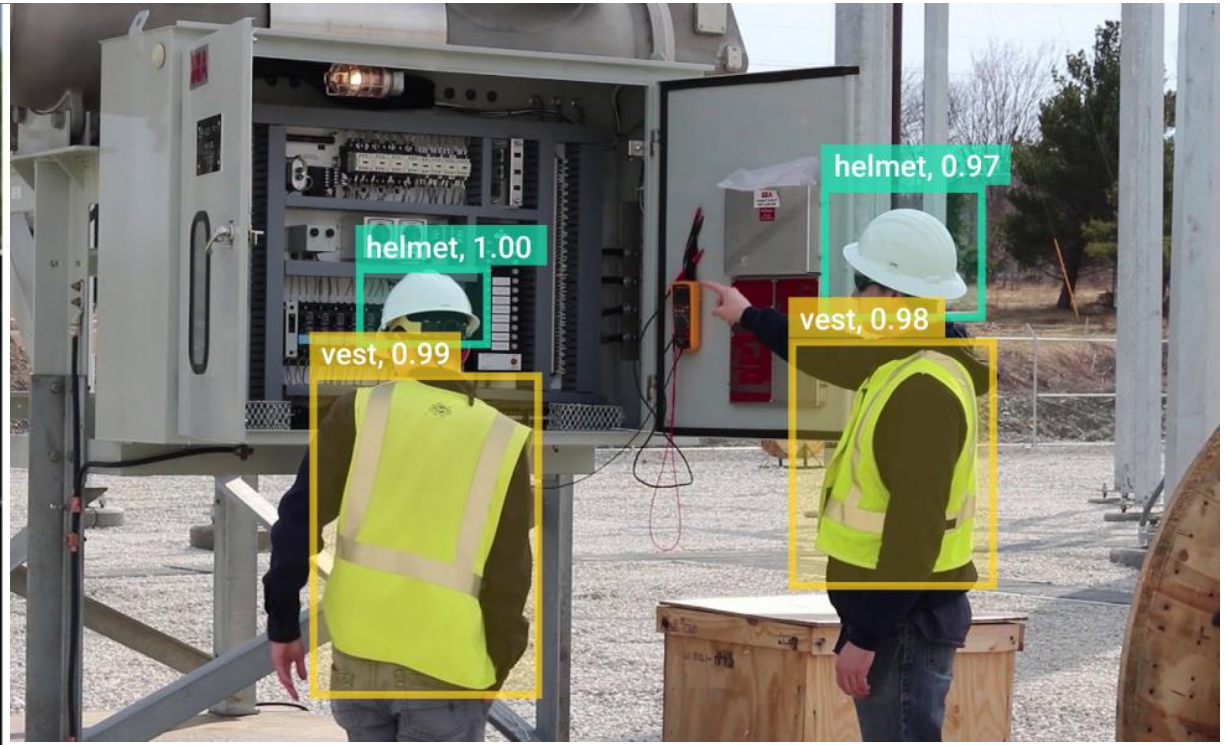


Image Segmentation



Object Detection and Tracking

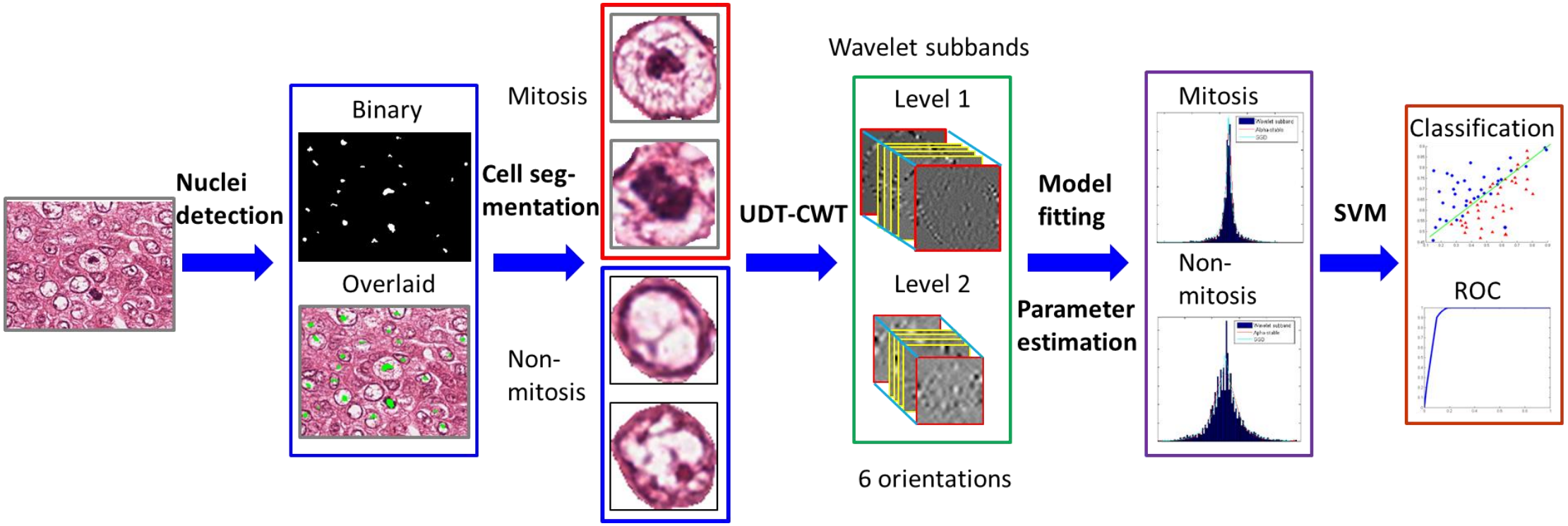


Gesture Recognition

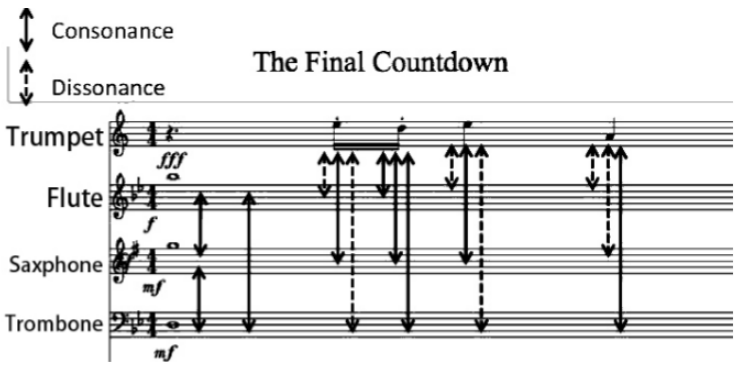
Represent pose as a
set of 14 joint positions:

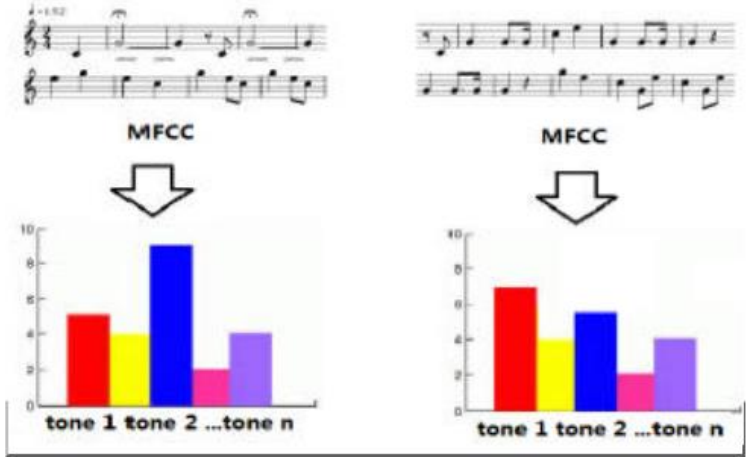
Left / right foot
Left / right knee
Left / right hip
Left / right shoulder
Left / right elbow
Left / right hand
Neck
Head top

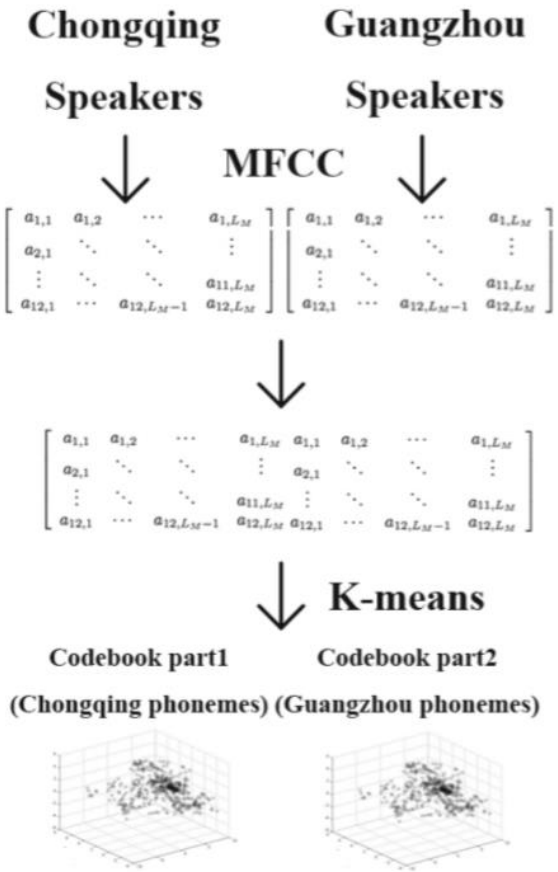
Medical Image Analysis

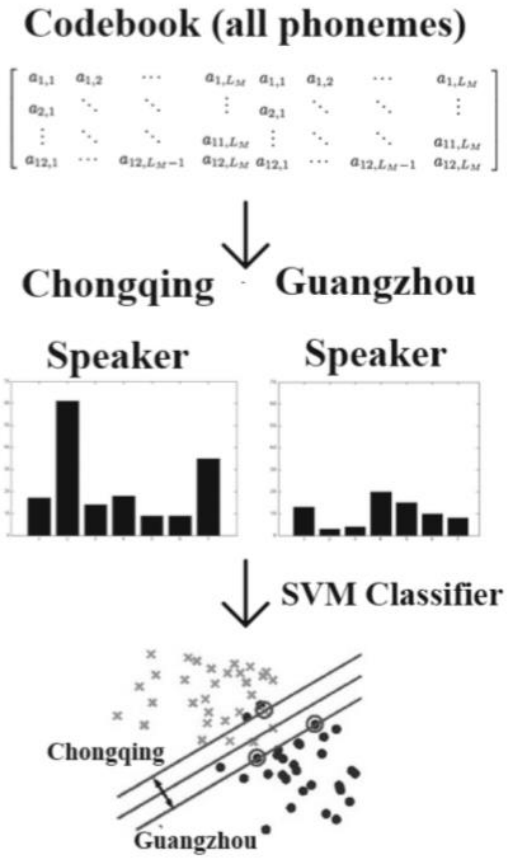


Music and Sound

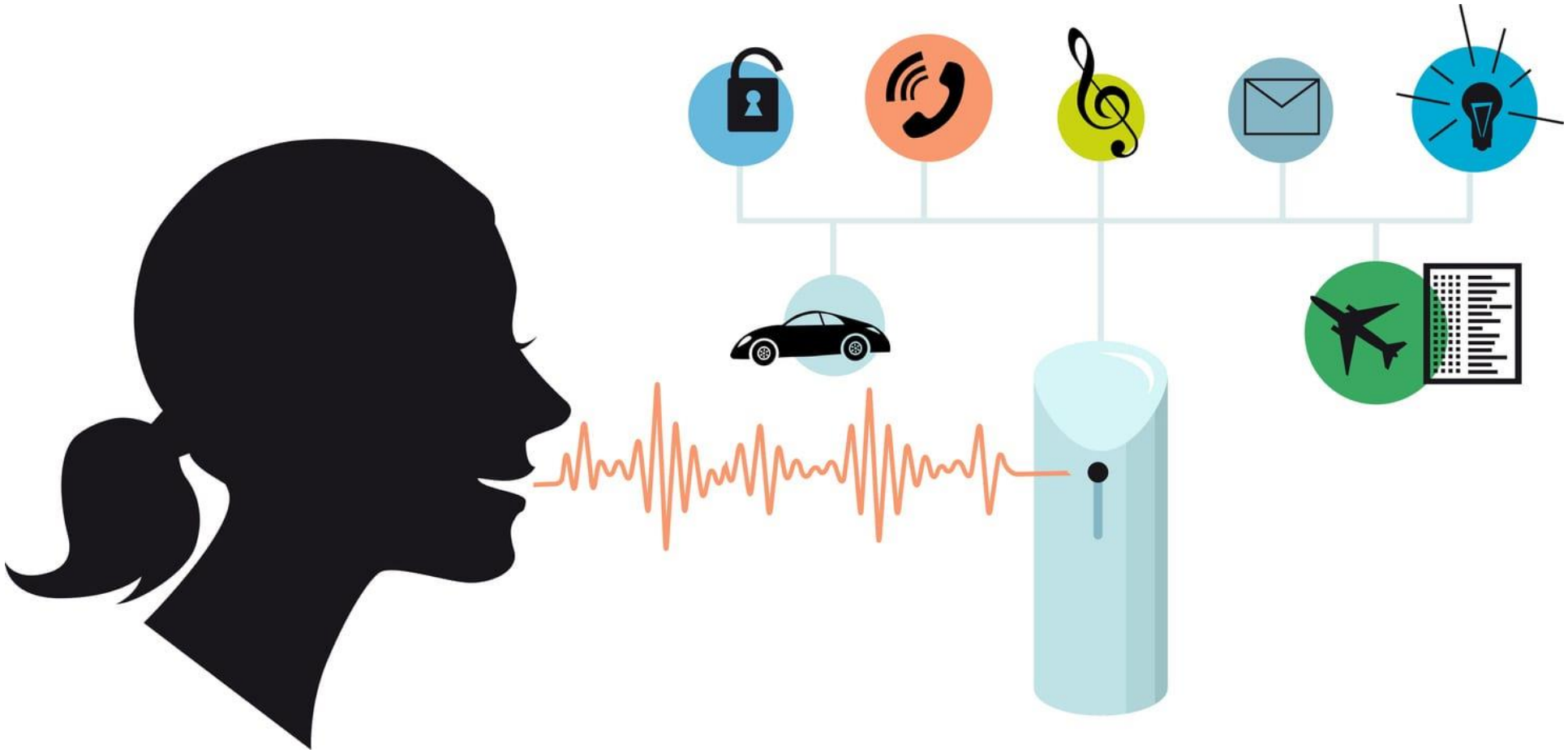




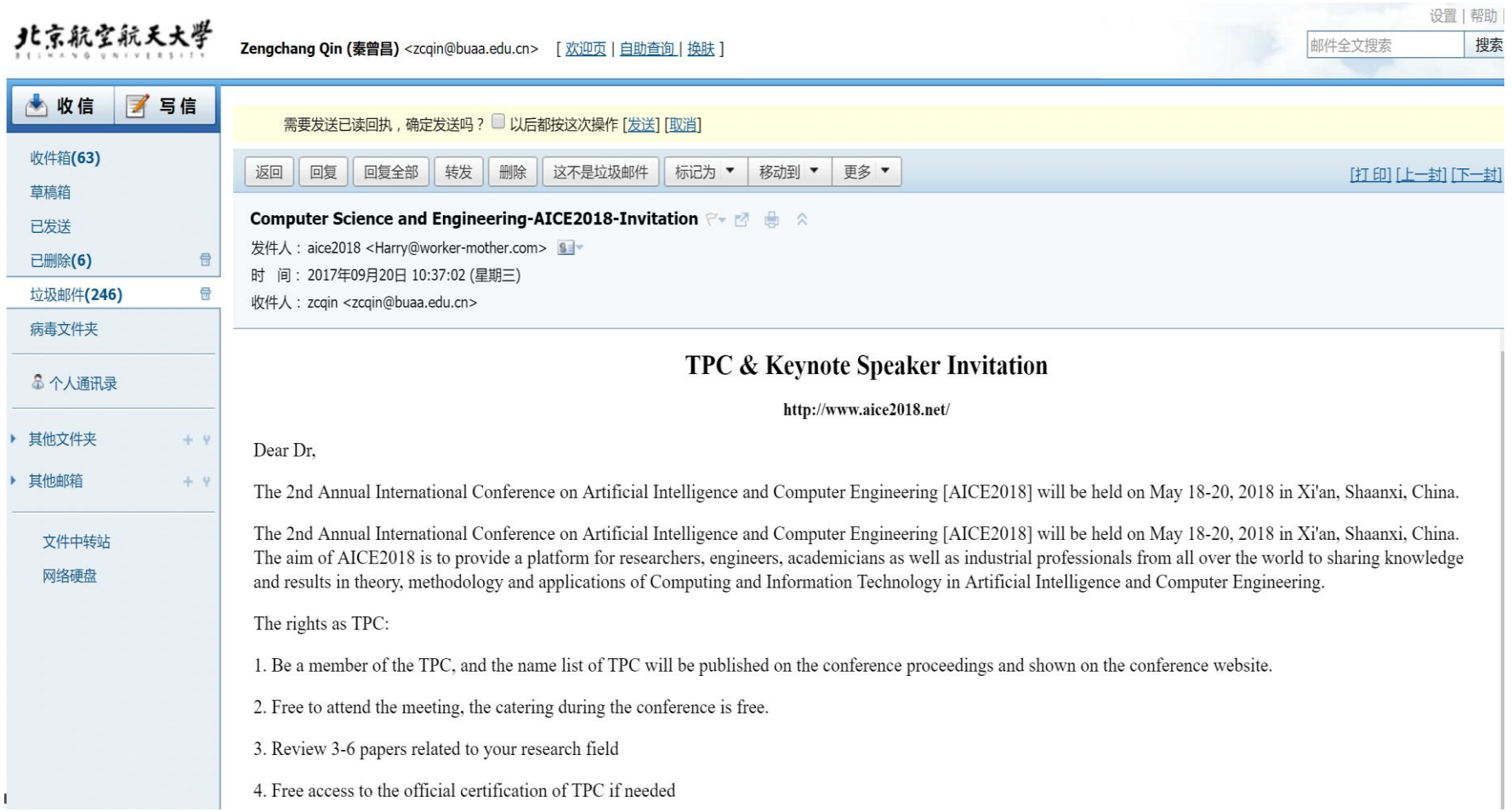




Speech Recognition



Spam Emails



北京航空航天大学 BEIHANG UNIVERSITY

Zengchang Qin (秦曾昌) <zcqin@buaa.edu.cn> [欢迎页 | 自助查询 | 换肤]

设置 | 帮助 | 邮件全文搜索 搜索

收信 写信

收件箱(63)
草稿箱
已发送
已删除(6)
垃圾邮件(246)
病毒文件夹
个人通讯录
其他文件夹
其他邮箱
文件中转站
网络硬盘

需要发送已读回执，确定发送吗？ 以后都按这次操作 [发送] [取消]

返回 回复 回复全部 转发 删除 这不是垃圾邮件 标记为 移动到 更多 [打印] [上一封] [下一封]

Computer Science and Engineering-AICE2018-Invitation

发件人： aice2018 <Harry@worker-mother.com>
时　间： 2017年09月20日 10:37:02 (星期三)
收件人： zcqin <zcqin@buaa.edu.cn>

TPC & Keynote Speaker Invitation

http://www.aice2018.net/

Dear Dr,

The 2nd Annual International Conference on Artificial Intelligence and Computer Engineering [AICE2018] will be held on May 18-20, 2018 in Xi'an, Shaanxi, China.

The 2nd Annual International Conference on Artificial Intelligence and Computer Engineering [AICE2018] will be held on May 18-20, 2018 in Xi'an, Shaanxi, China. The aim of AICE2018 is to provide a platform for researchers, engineers, academicians as well as industrial professionals from all over the world to sharing knowledge and results in theory, methodology and applications of Computing and Information Technology in Artificial Intelligence and Computer Engineering.

The rights as TPC:

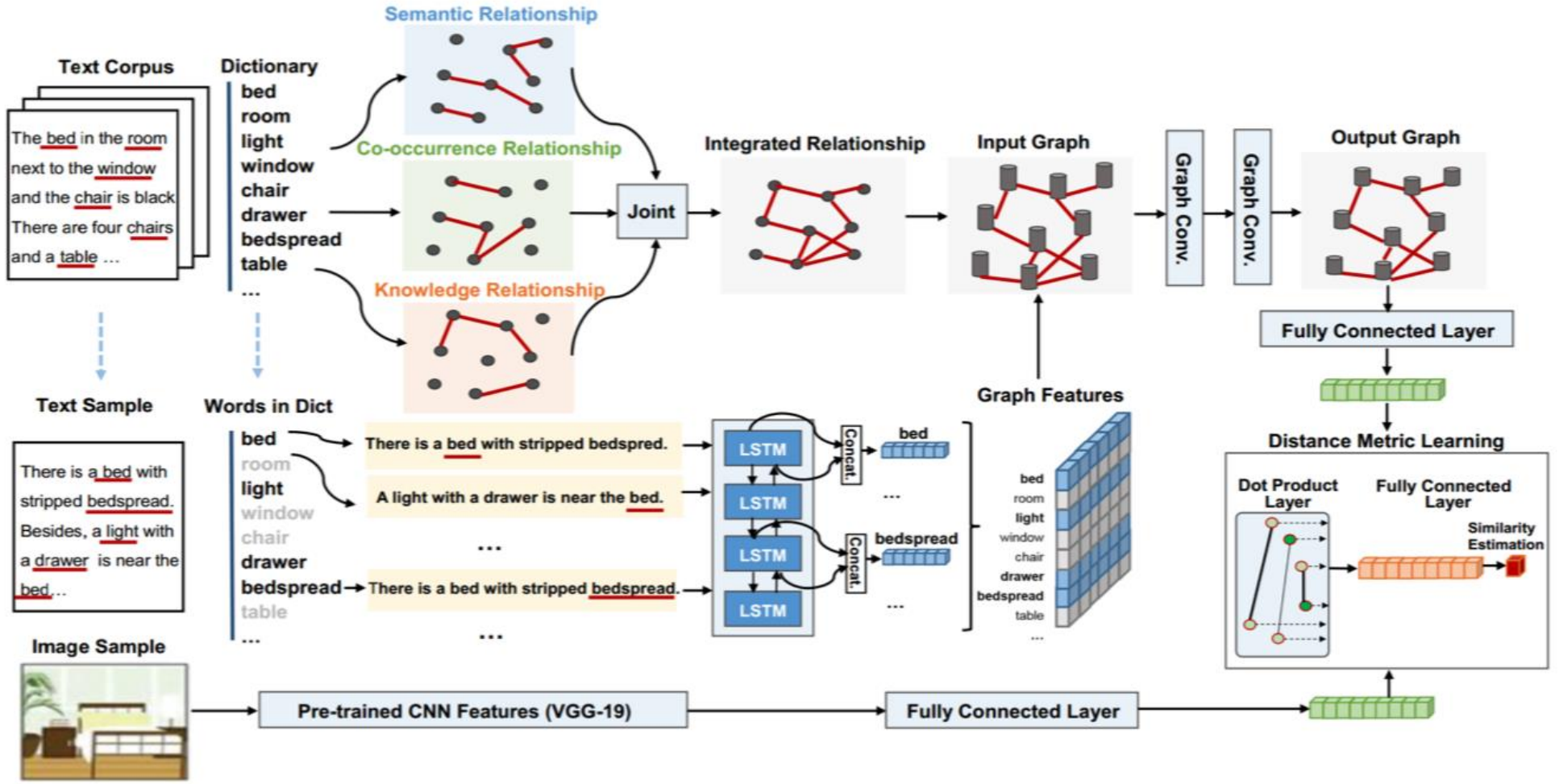
1. Be a member of the TPC, and the name list of TPC will be published on the conference proceedings and shown on the conference website.

2. Free to attend the meeting, the catering during the conference is free.

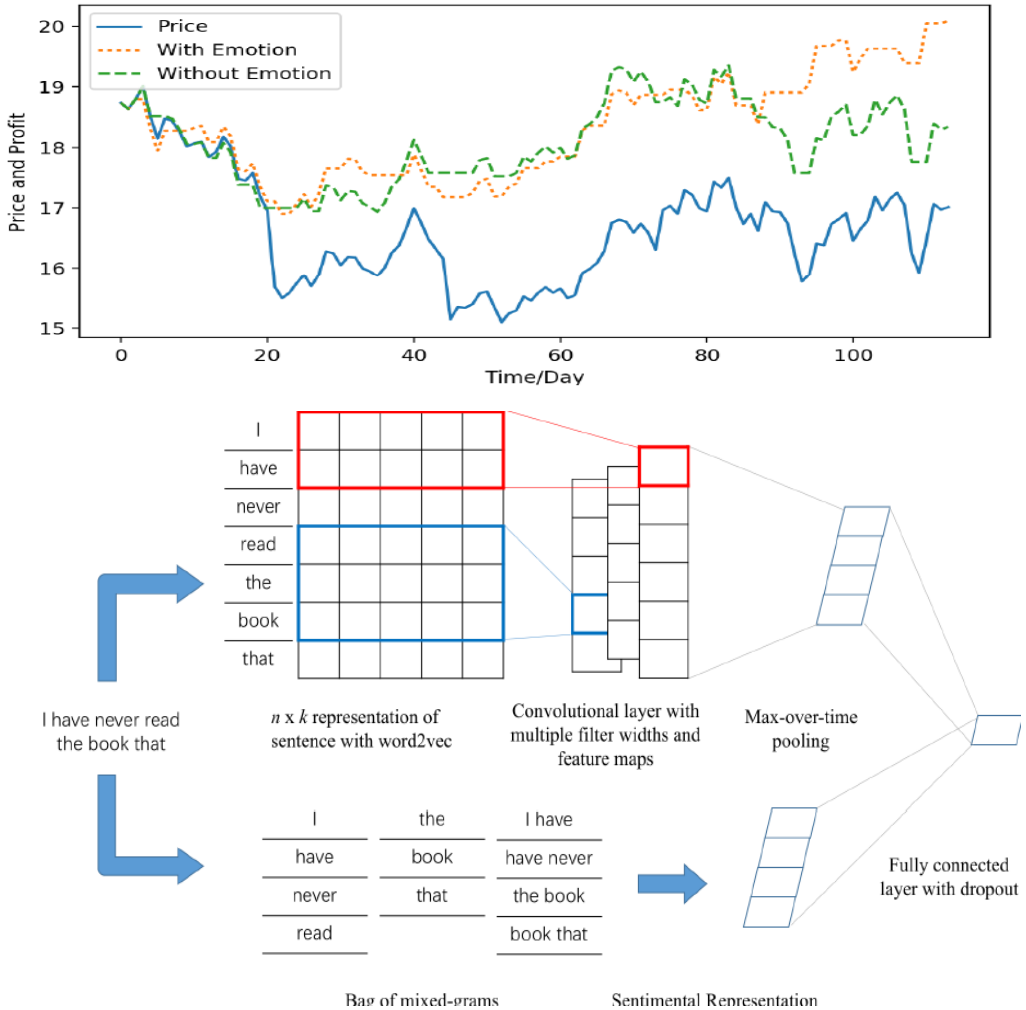
3. Review 3-6 papers related to your research field

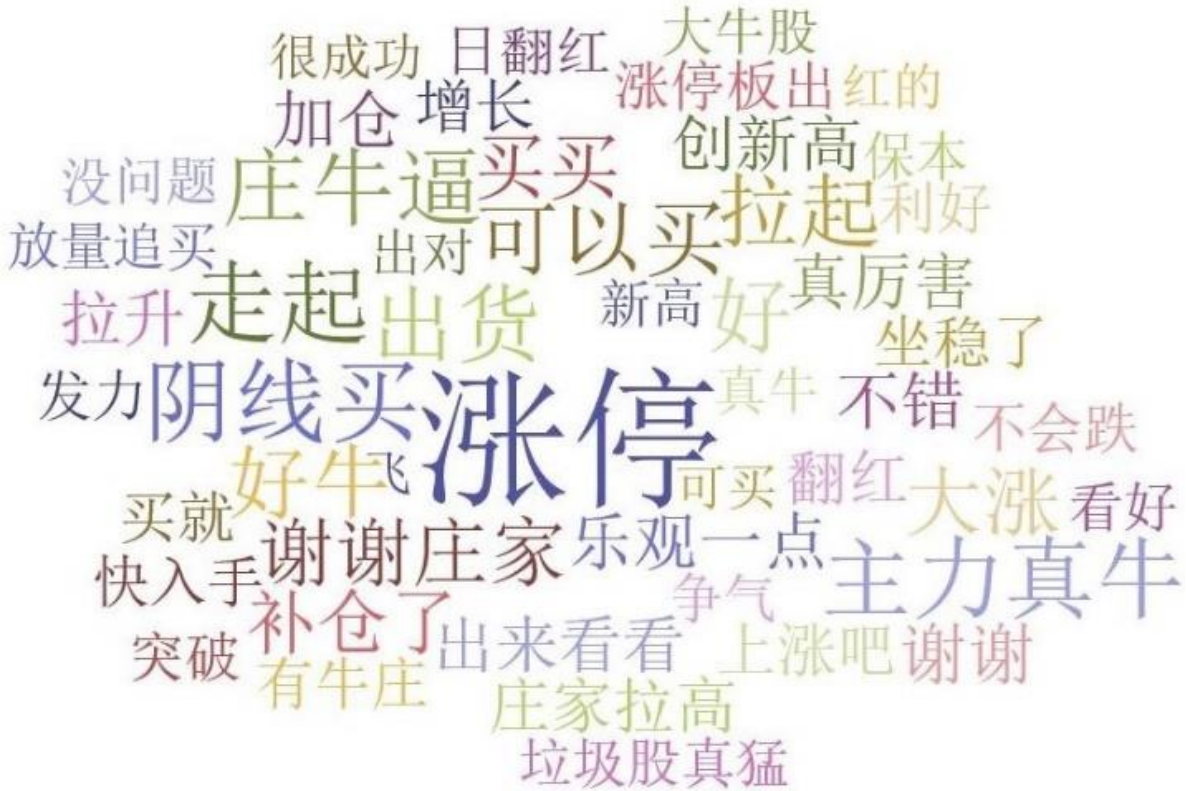
4. Free access to the official certification of TPC if needed

Cross-Modal Retrieval

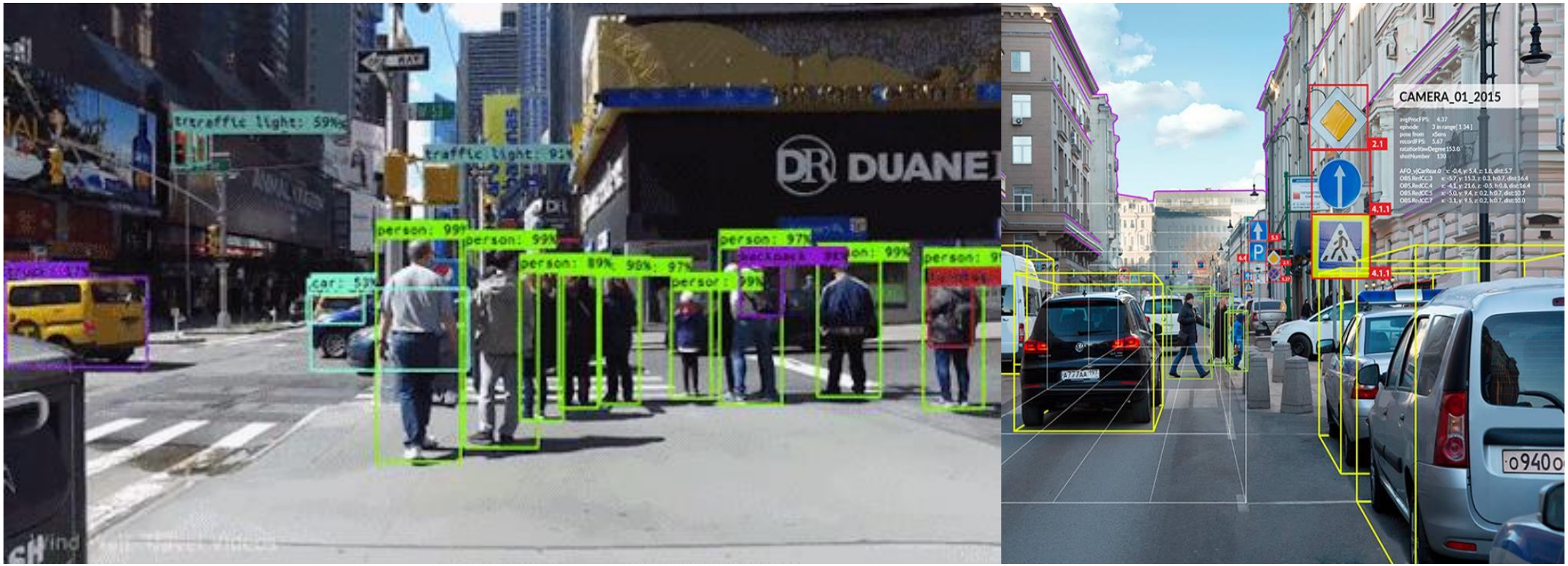


Financial Data Analysis

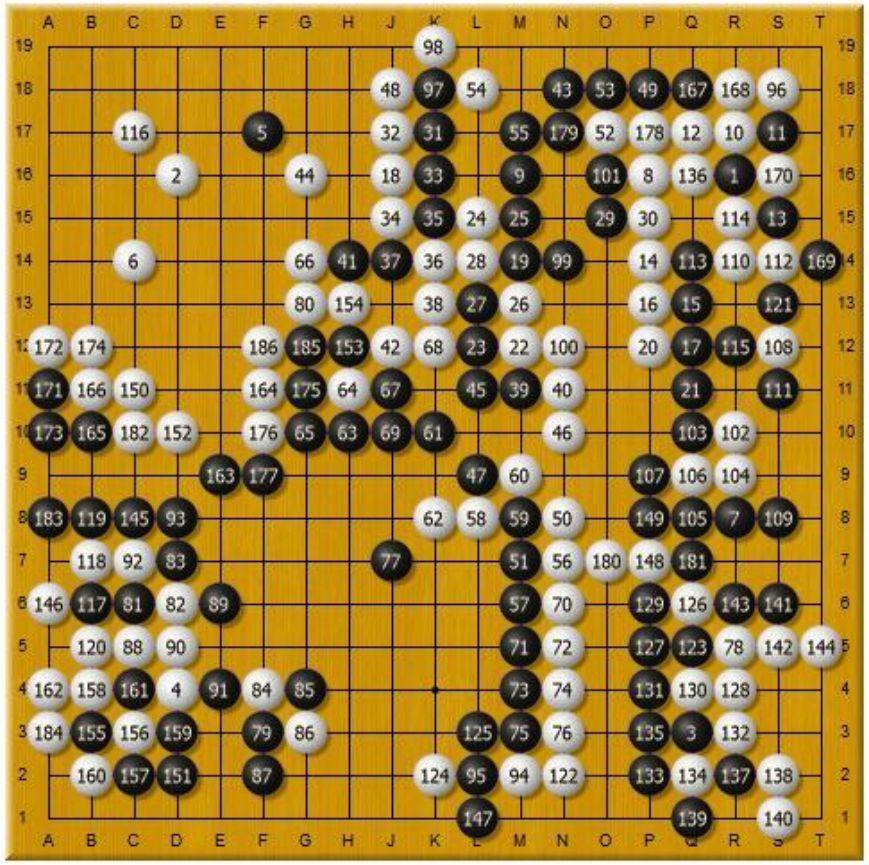




Automatic Driving

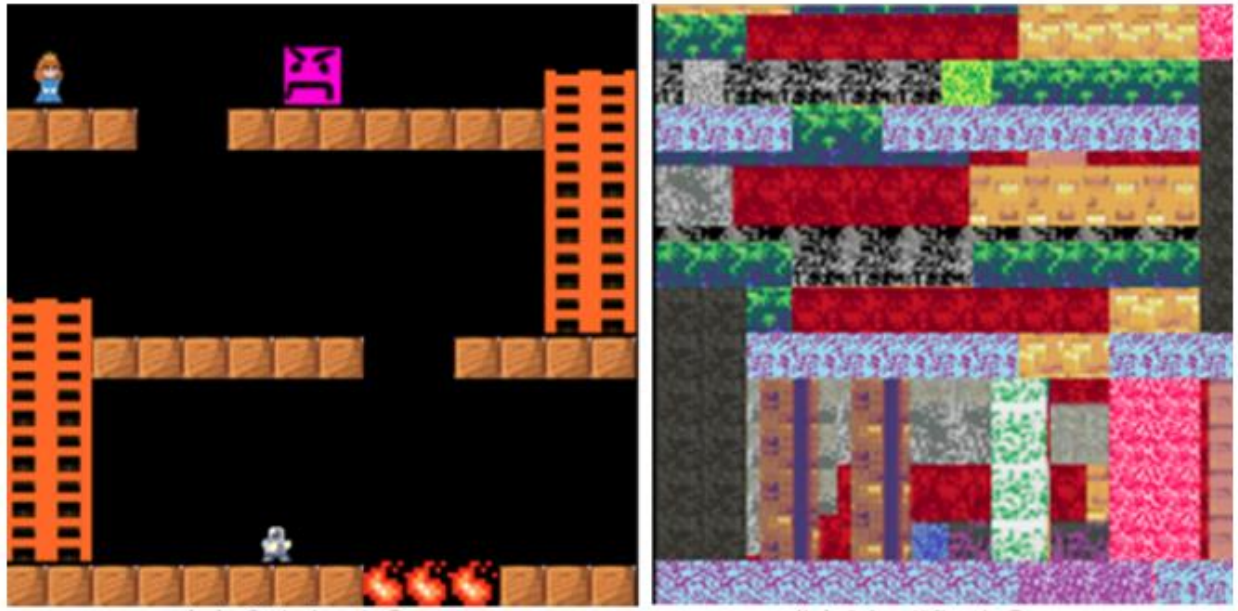


AlphaGo




Google DeepMind
Challenge Match
8 - 15 March 2016
AlphaGo
AlphaGo
Lee Sedol

Learning without Human Priors



(a) Original Game (b) Modified Game

https://openreview.net/pdf?id=Hk91SGWR- (Griffths, Berkeley)

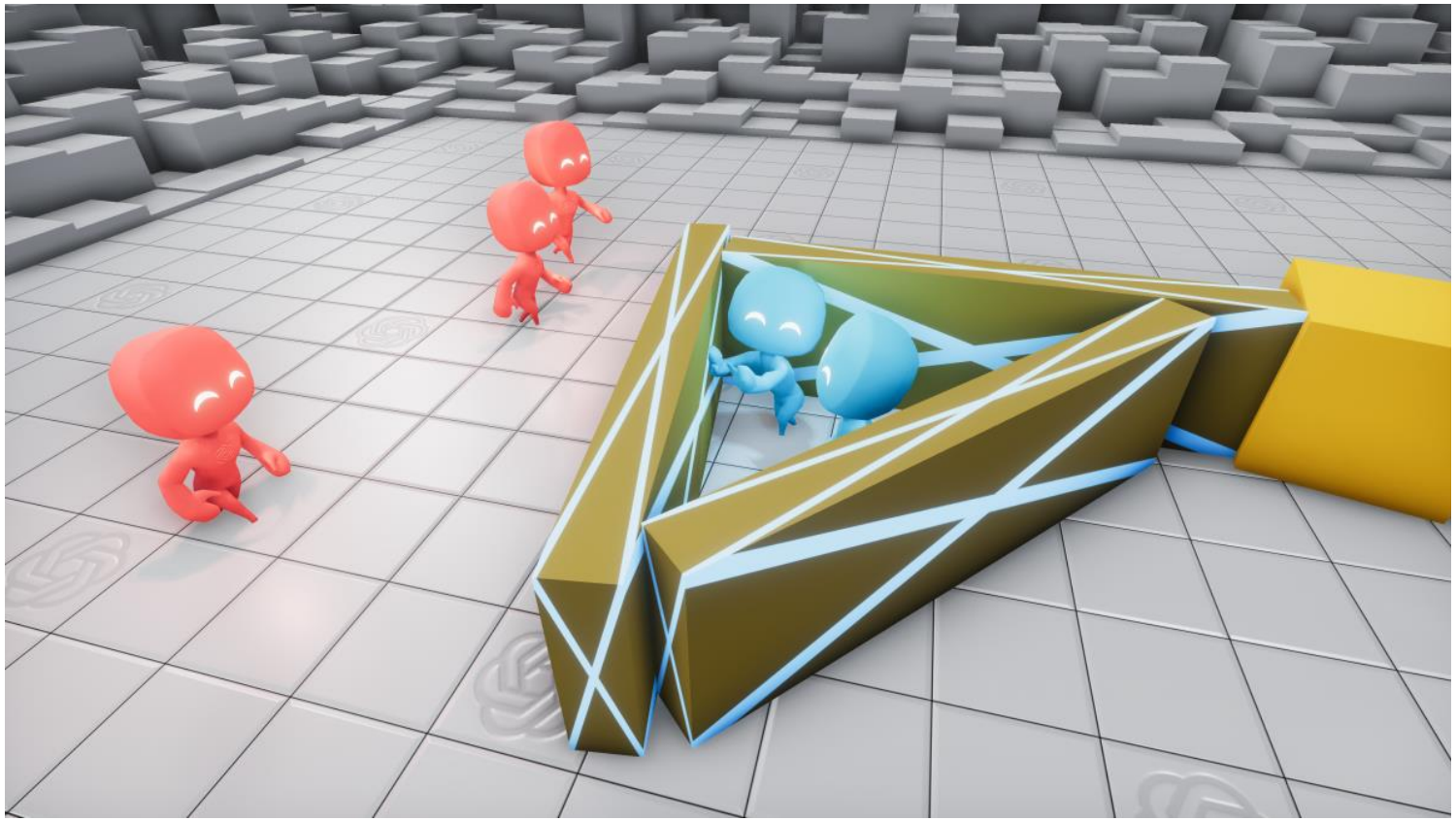
https://deepmind.com/blog/article/open-sourcing-psychlab

Arcade games were modified to re-render the game's textures. In the modified game, humans performed extremely poorly. In contrast, a Deep Learning system performed equivalently for both games. Deep Learning systems do not need to use human priors. On the flip side, a human can learn a game with fewer trials because we can exploit the use of existing human priors (or affordances). What this should tell you is that humans learn quickly by using our existing priors.

Hide and Seek

OpenAI Tried to Train AI Agents to Play Hide-And-Seek but Instead They Were Shocked by What They Learned.

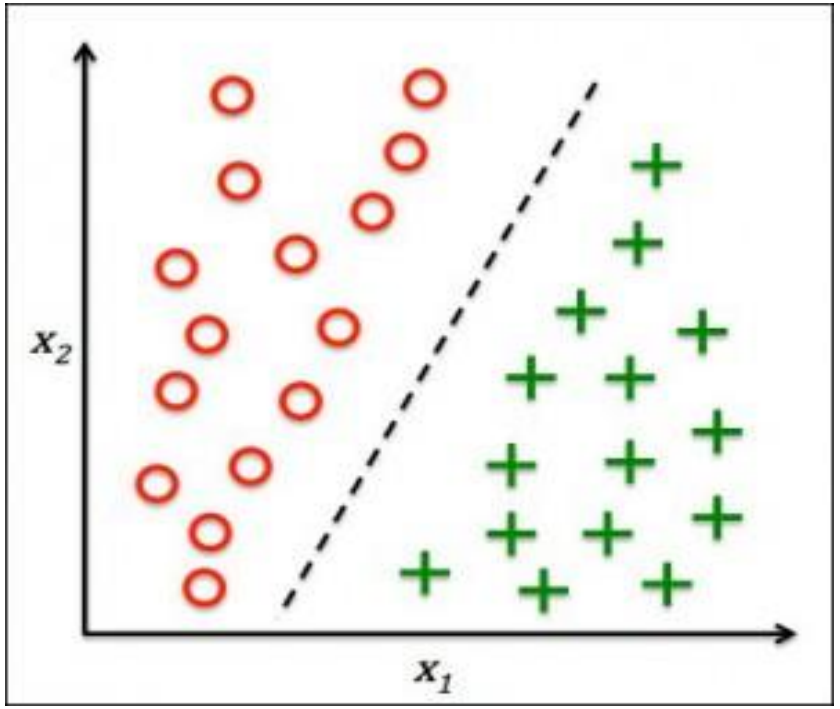
Learning by competition is one of the emerging paradigms in AI that fatefully resembles how our knowledge evolves as human species.

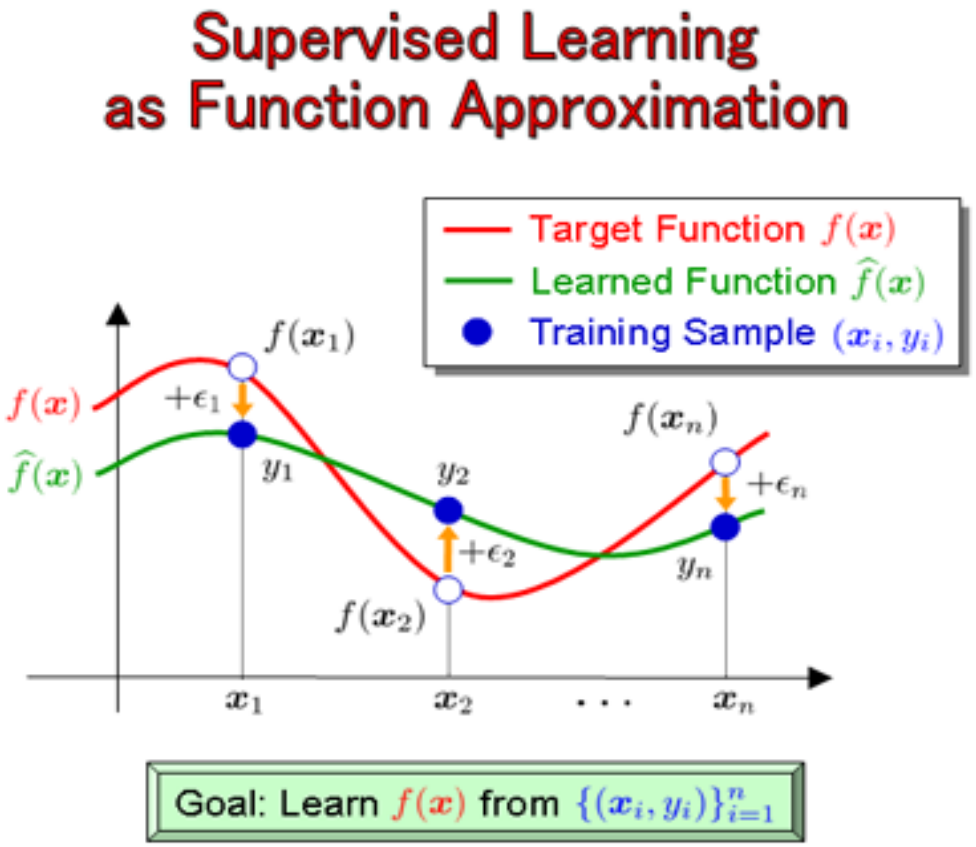


https://www.kdnuggets.com/2019/10/openai-tried-train-ai-agents-play-hide-seek-instead-shocked-learned.html

Supervised Learning

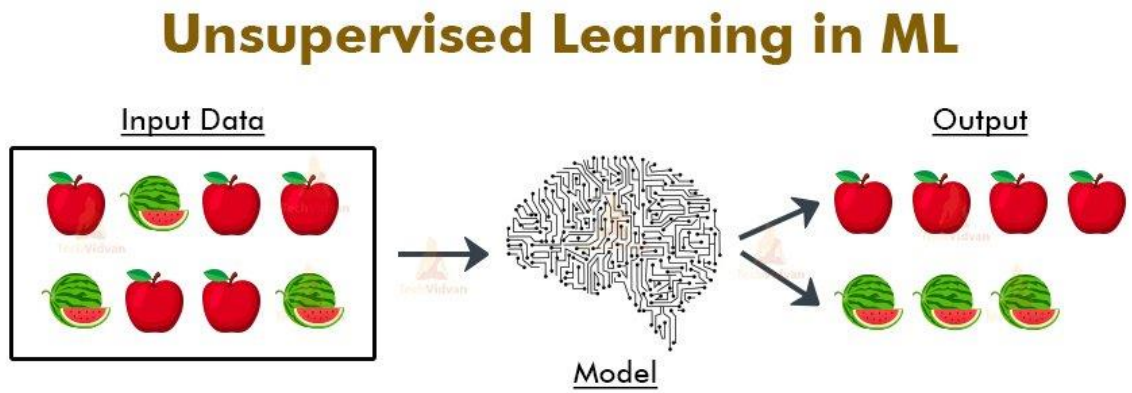
Classification

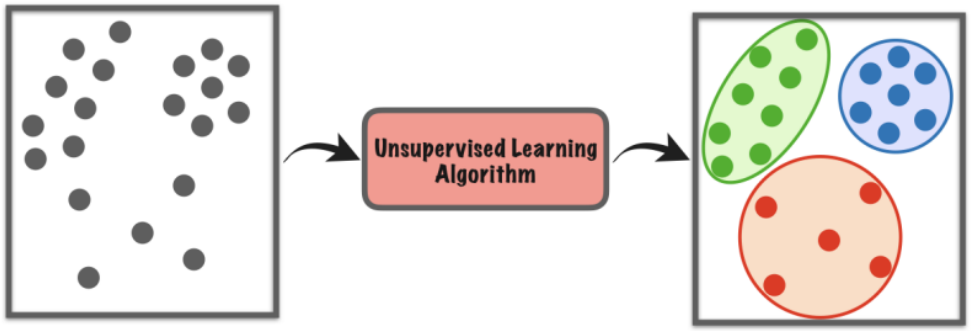


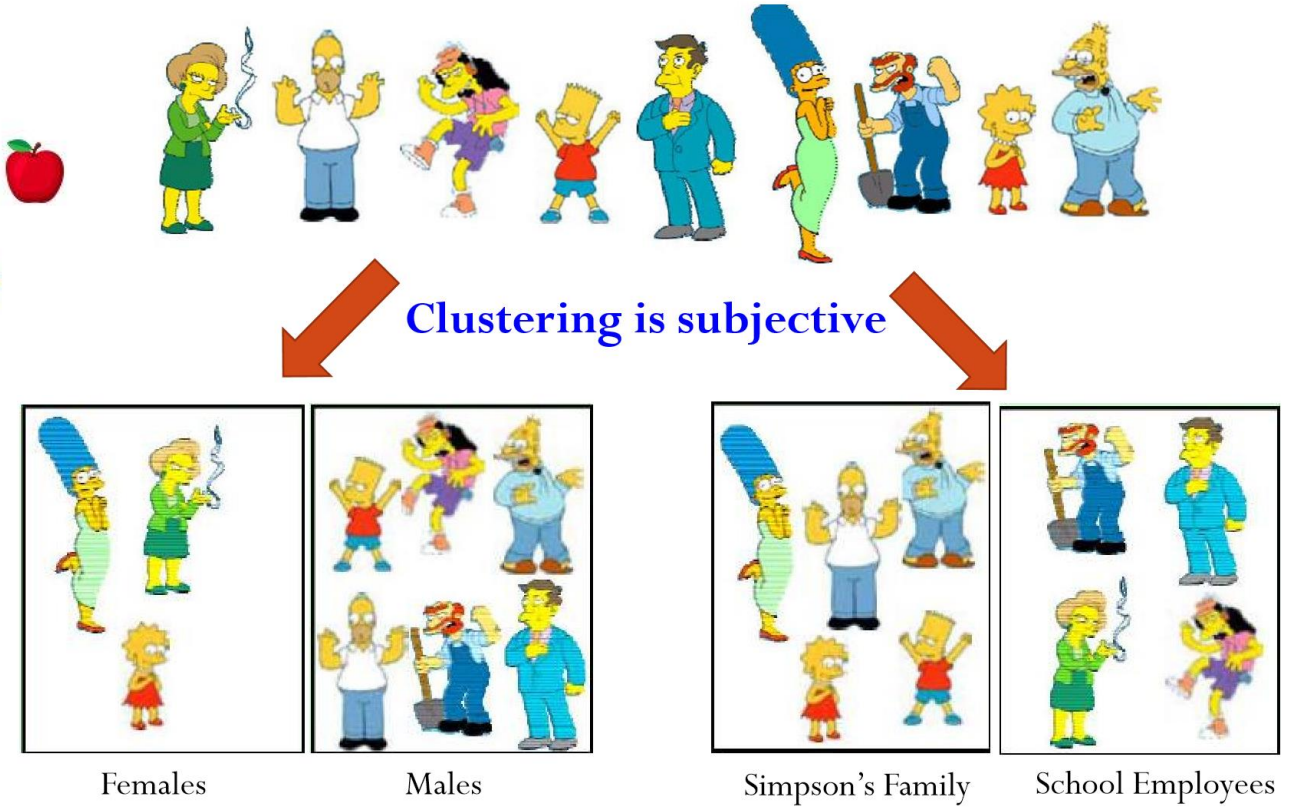


Classification is also a prediction of f(x) -> {0, 1}

Unsupervised Learning



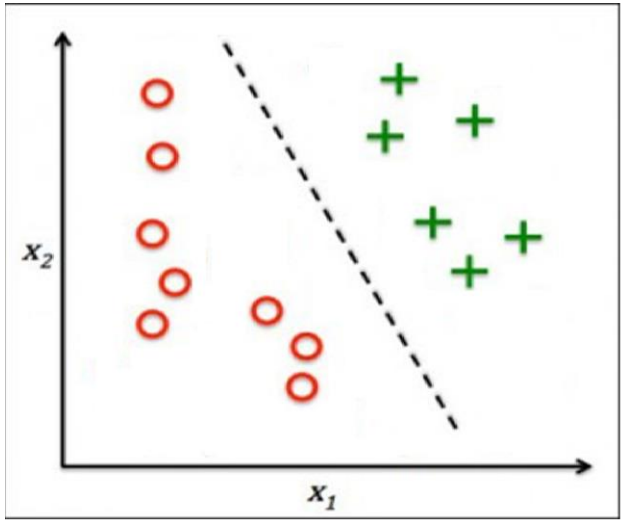




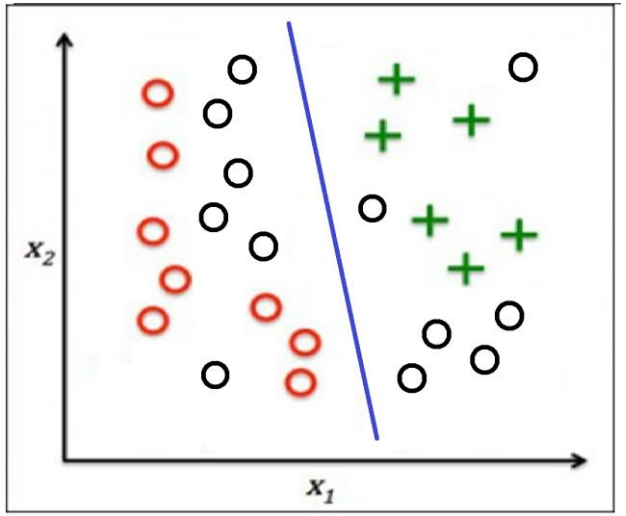
No target labels are given, similarity among data needs to be measured quantitatively.

Semi-Supervised Learning

- All labeled data

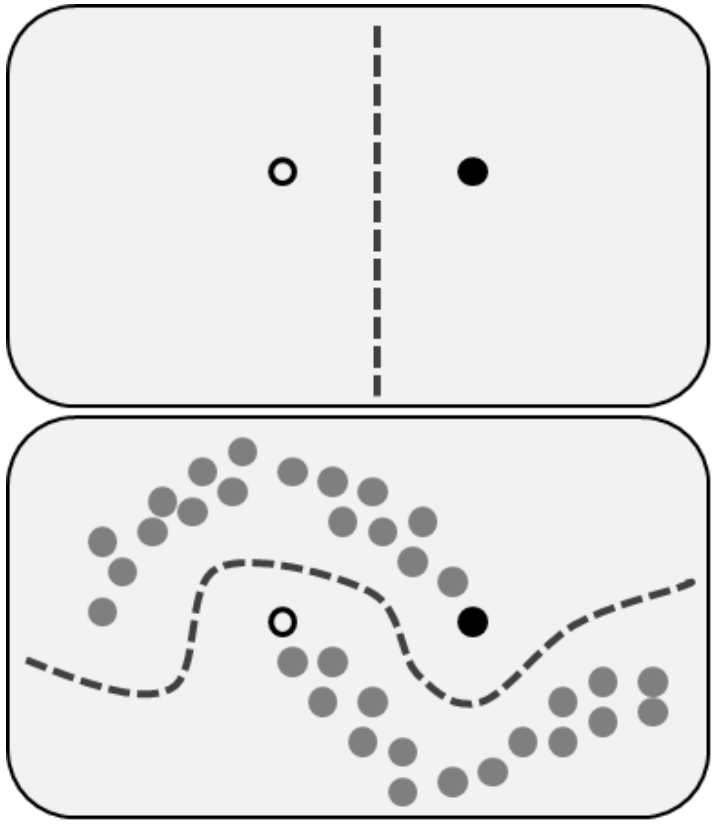


- With unlabeled data



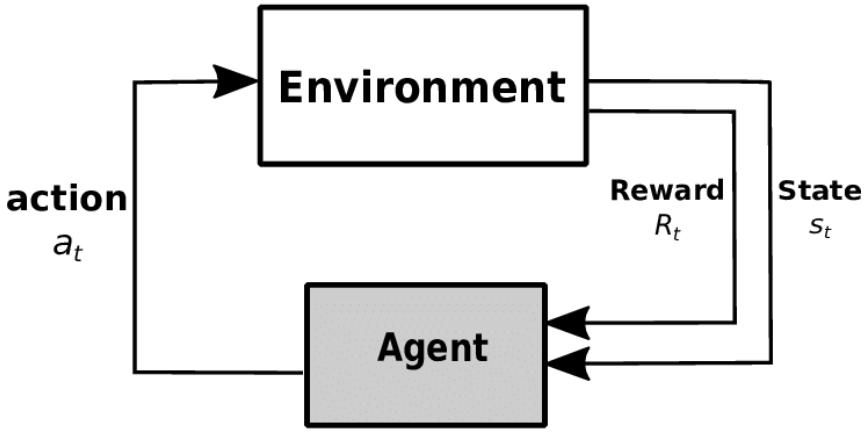
1. Points that are close to each other are more likely to ***share a label****.*

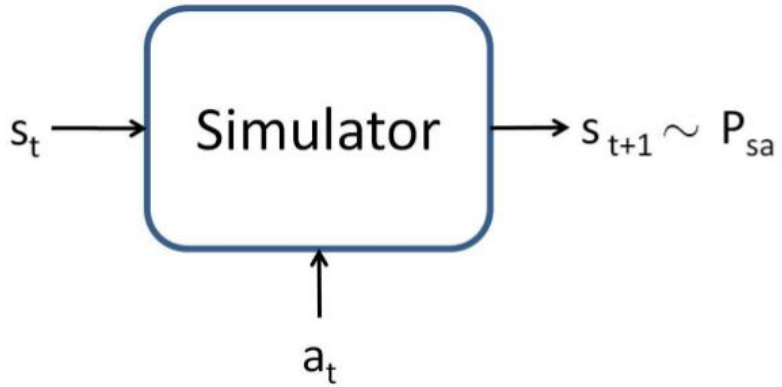
2. The data lie approximately on a ***manifold*** *of much lower dimension than the input space*



Data labelling is expensive!

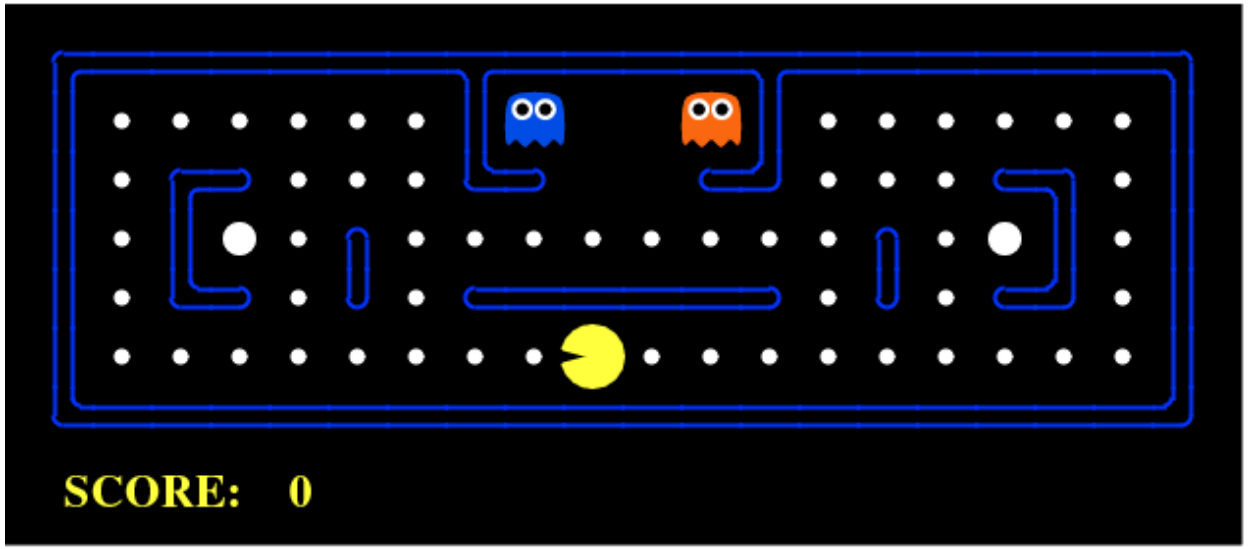
Reinforcement Learning





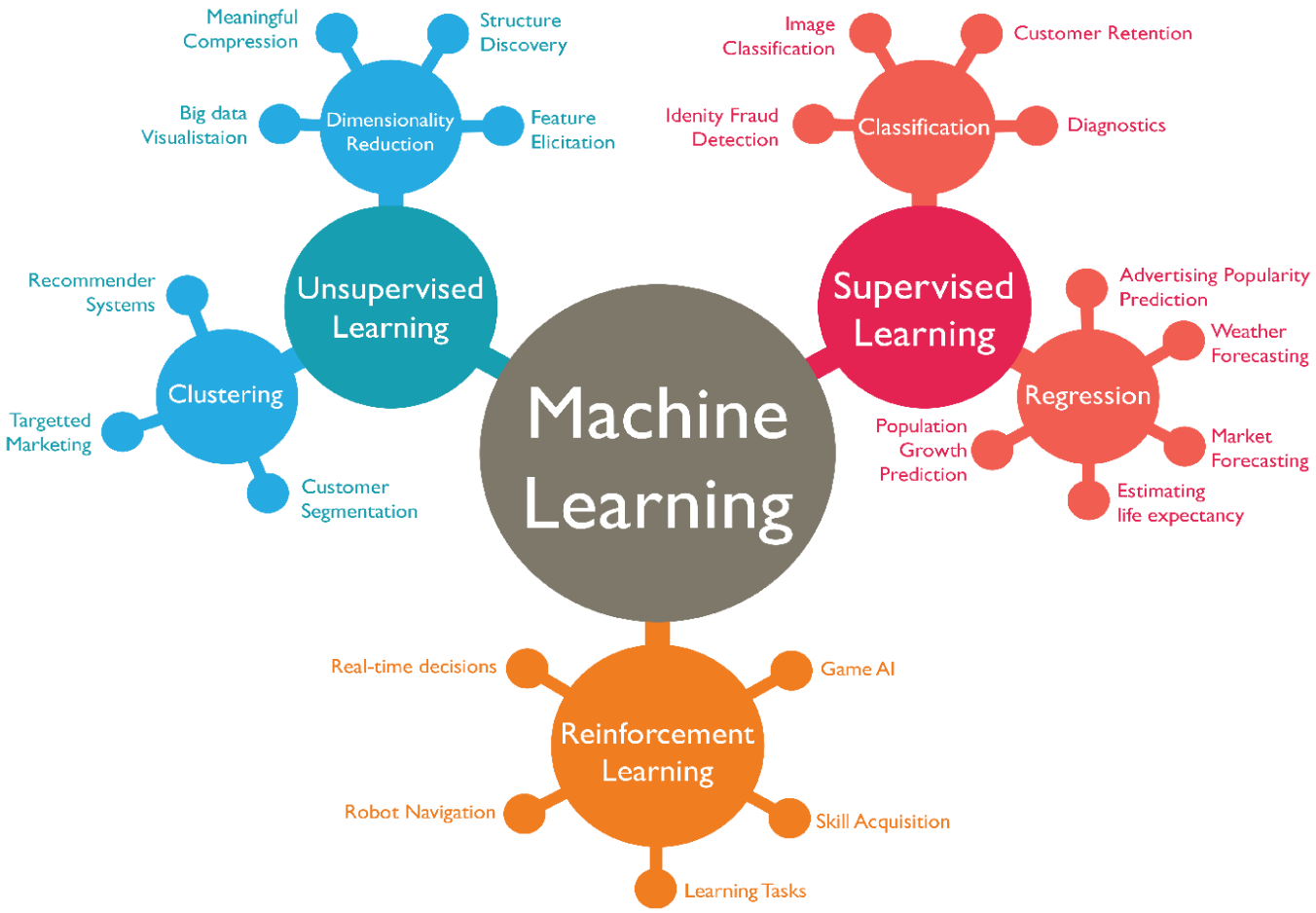
Q: Advantages and disadvantages of RL?

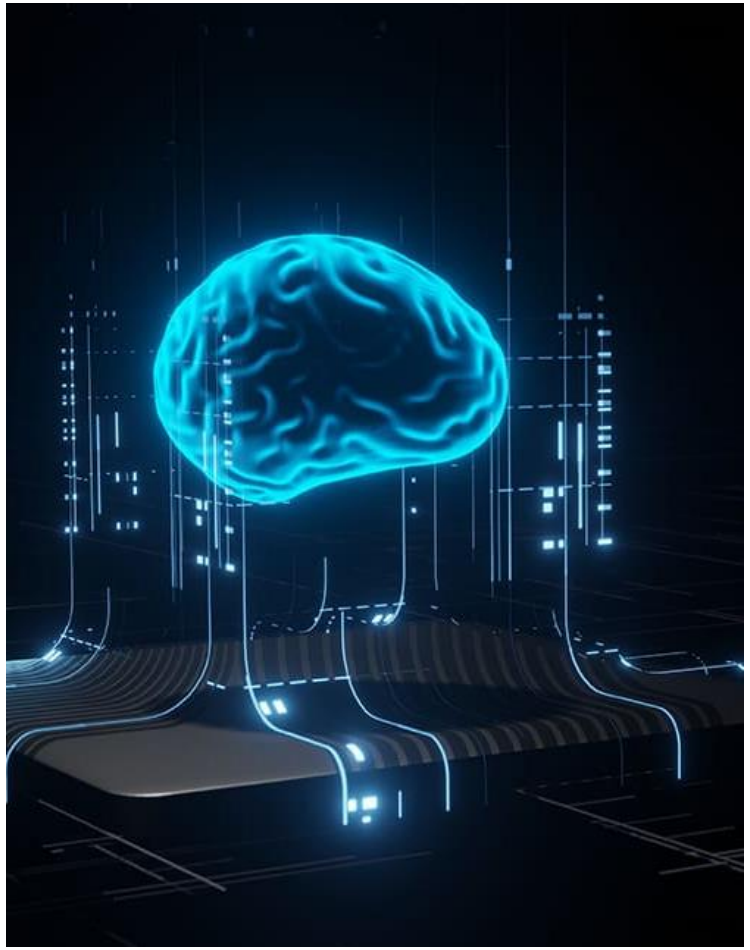
In order to build an optimal policy, the agent faces the dilemma of exploring new states while maximizing its overall reward at the same time. This is called **Exploration vs Exploitation** trade-off. To balance both, the best overall strategy may involve short term sacrifices.



- PAC Man with Reinforcement Learning

Paradigm of Machine Learning





References

[1]https://med.libretexts.org/Bookshelves/Pharmacology_and_Neuroscience/Book%3A_Computational_Cognitive_Neuroscience_(O%27Reilly_and_Munakata)/6%3A_Preception_and_Attention/6.3%3A_Oriented_Edge_Detectors_in_Primary_Visual_Cortex

[2] https://people.eecs.berkeley.edu/~malik/papers/SM-ncut.pdf