

CSI 436/536 (Fall 2024)

Machine Learning

Lecture 20: Advanced Topic: Decision Making

Chong Liu

Assistant Professor of Computer Science

Nov 21, 2024

Machine learning overview

- Supervised learning
 - Tasks: classification, regression
 - Techniques: regularization, error decomposition, kernel, ensemble, neural network
- Unsupervised learning
 - Clustering, dimension reduction
- Some many advanced topics:
 - More tasks: ranking
 - More theoretical understanding: statistical learning theory
 - More settings: reinforcement learning
 - More applications: computer vision, natural language processing, speech recognition
 - More perspective: rule learning
 - More trustworthy: explainable ML, privacy / fairness issues in ML
 - ...

Today

- Decision making-based machine learning
 - Active learning
 - Disagreement-based algorithm
 - Confidence-based algorithm
 - Clustering-based algorithm
 - Bandits
 - Reinforcement learning

Today

- Decision making in machine learning
 - Active learning
 - Disagreement-based algorithm
 - Confidence-based algorithm
 - Clustering-based algorithm
 - Bandits
 - Reinforcement learning

Binary image classification tasks



Dog or mop?

Binary image classification tasks



Dog or croissant?

Binary image classification tasks



Dog or bagel?

Binary image classification tasks



Dog or muffin?

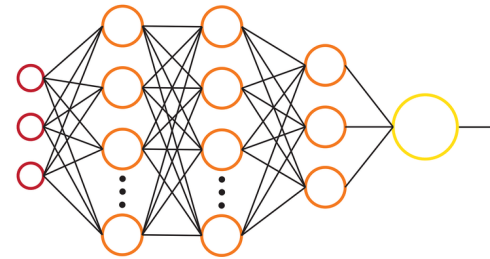
Binary image classification tasks



Dog or fried chicken?

Recap: Supervised learning

- How to achieve good performance?
 - Advanced techniques: kernel, ensemble, neural networks ...
 - Input **a lot of labeled images** of dogs or fried chicken

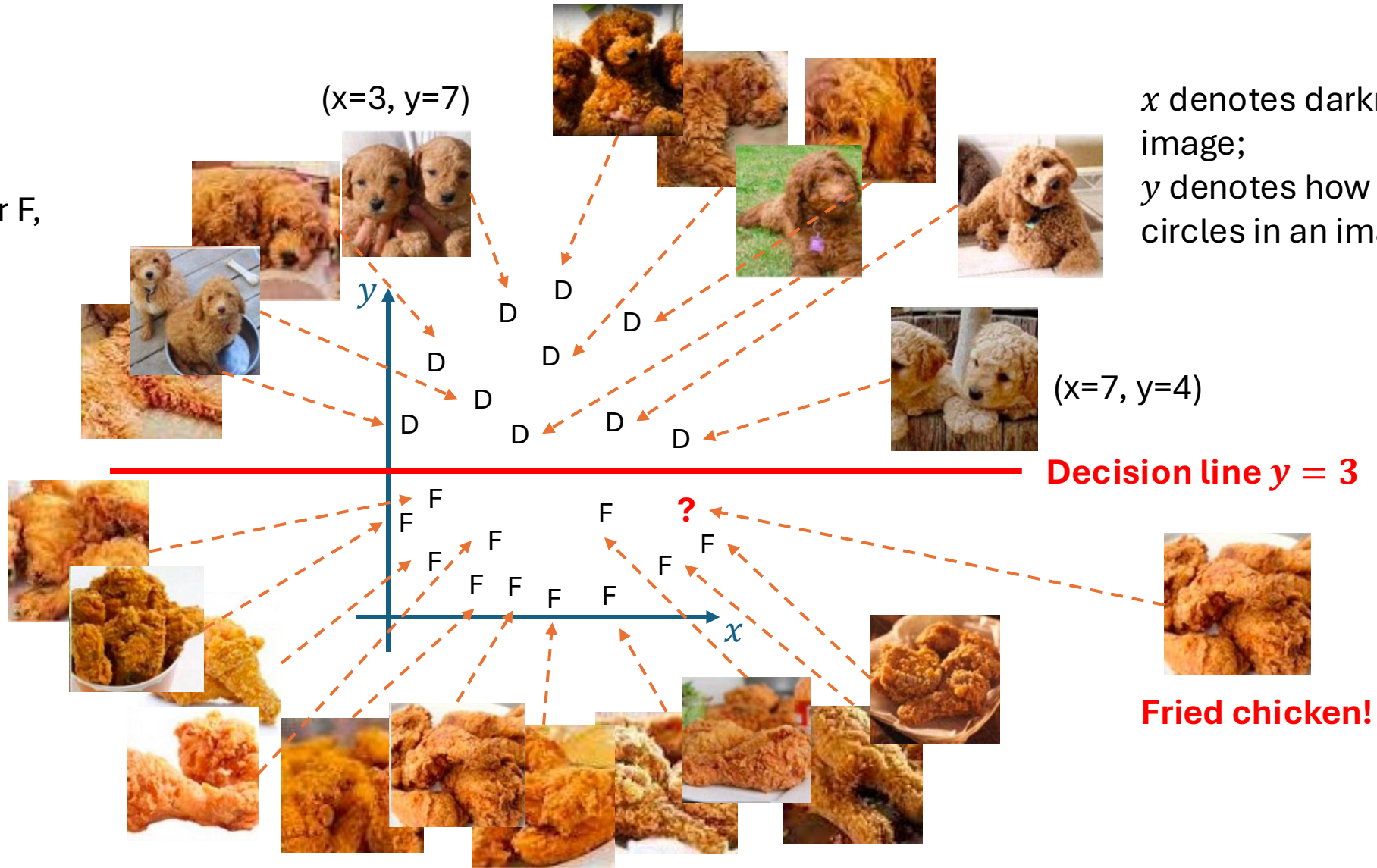


Given a new image, dog or fried chicken?



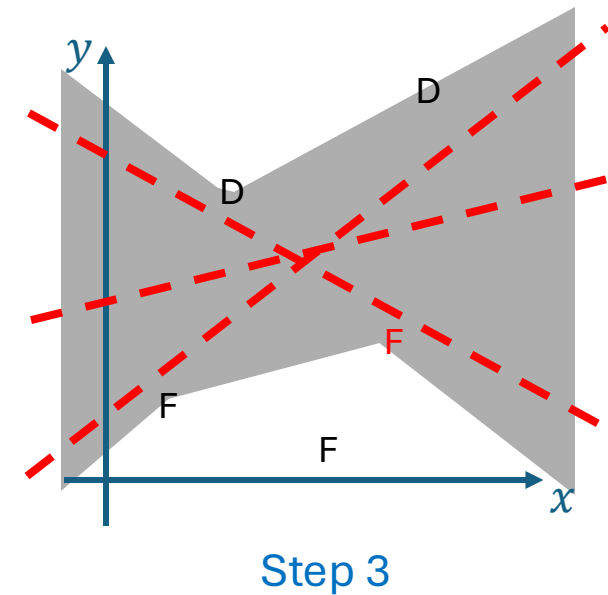
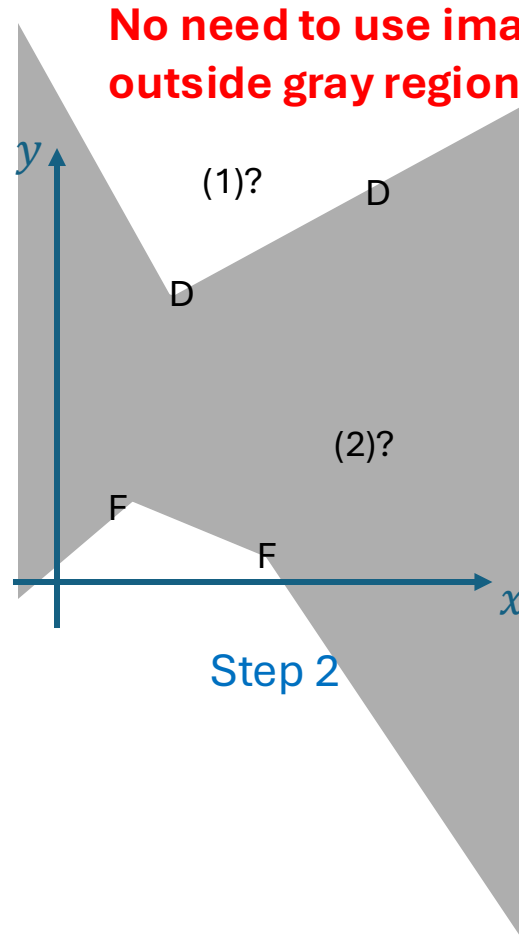
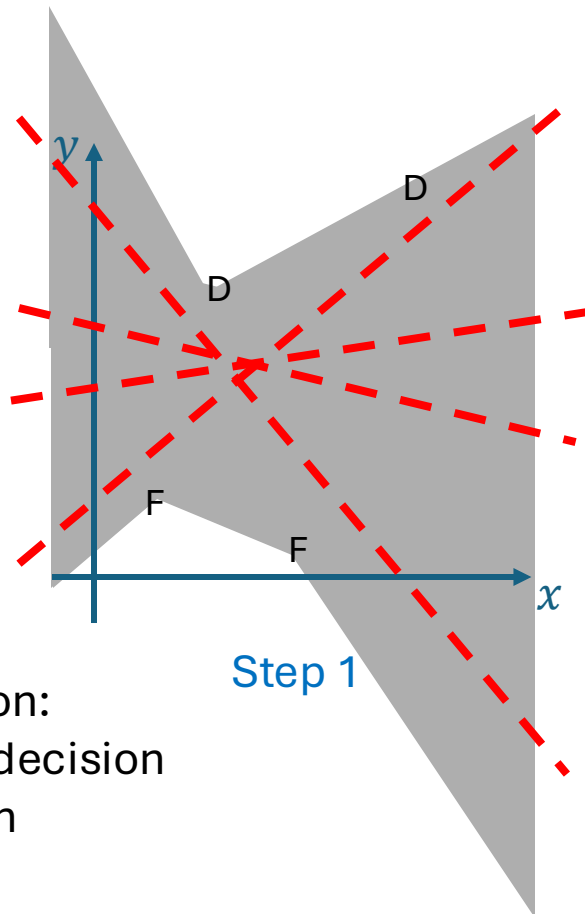
Binary classification in 2-dimension

Each point, D or F, represents an image.



Fewer images to learn the decision line?

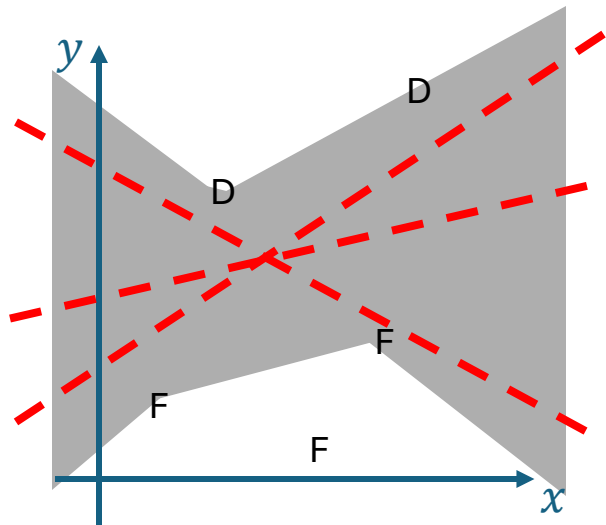
- Yes! Active learning selects important images.



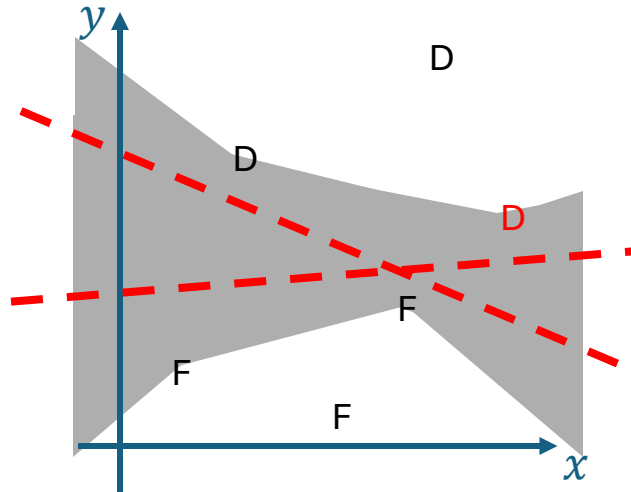
Gray region:
Feasible decision
line region

x denotes darkness of an image;
 y denotes how many black
circles in an image.

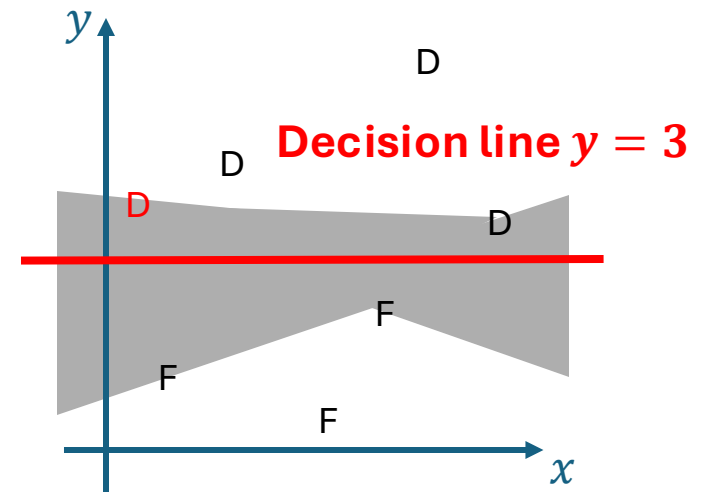
Moving forward ...



Step 3



Step 4



Step 5

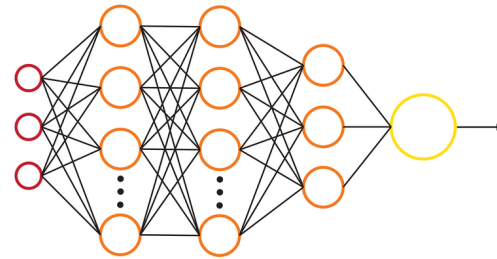
- Learn decision line with only 7 images!
 - Active learning only selects important images
 - images in gray region

Disagreement-based Active Learning

Confidence-based active learning



Labeled data

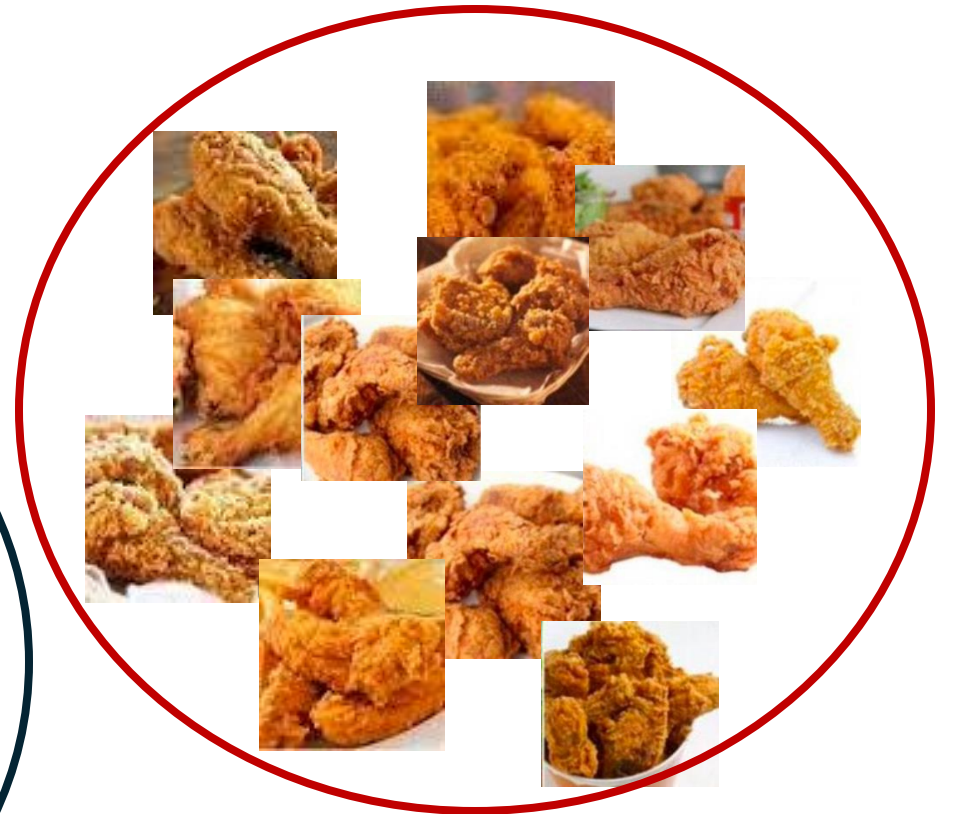
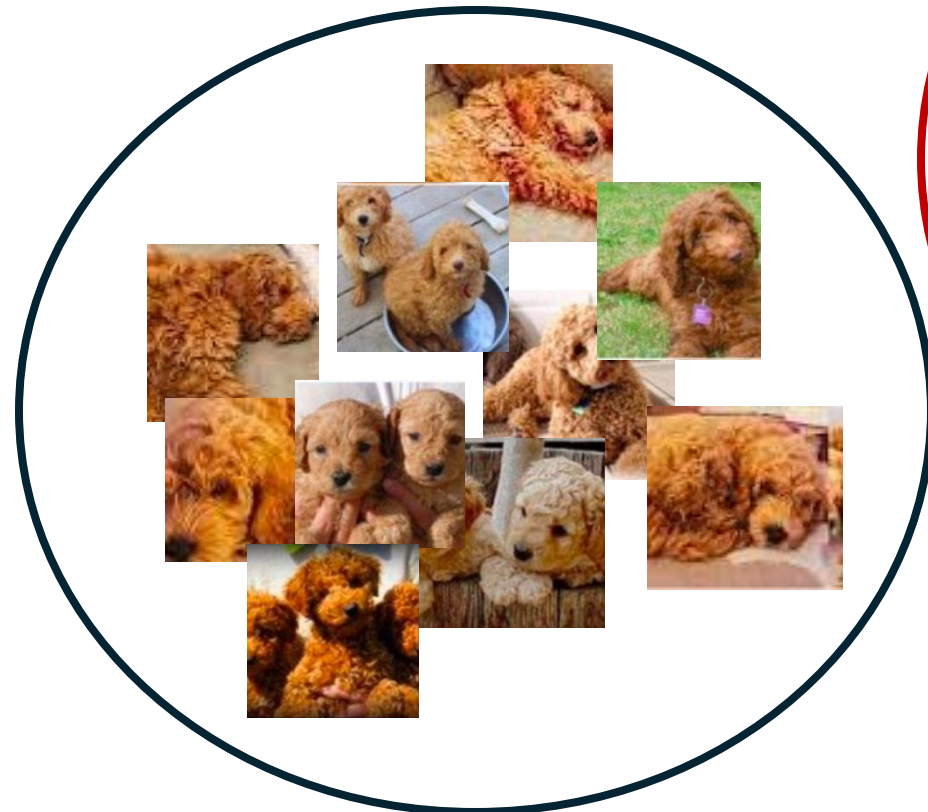


Unlabeled data

How to select the data point with least prediction confidence?

Clustering-based active learning

- Run k-means clustering
- Select samples from each cluster



Summary: active learning

- Framework:
 - Input: a small labeled dataset, a large unlabeled dataset
 - Output: a well-trained model
 - For $t = 1, 2, \dots$
 - Select a sample x_t from the unlabeled dataset
 - Get its label y_t
 - Update the model by adding (x_t, y_t)
- Goal:
 - Achieve good performance while **saving labeling cost** at the same time
- Can be super useful in practice!

Today

- Decision making in machine learning
 - Active learning
 - Disagreement-based algorithm
 - Confidence-based algorithm
 - Clustering-based algorithm
 - **Bandits**
 - Reinforcement learning

The real bandit machine

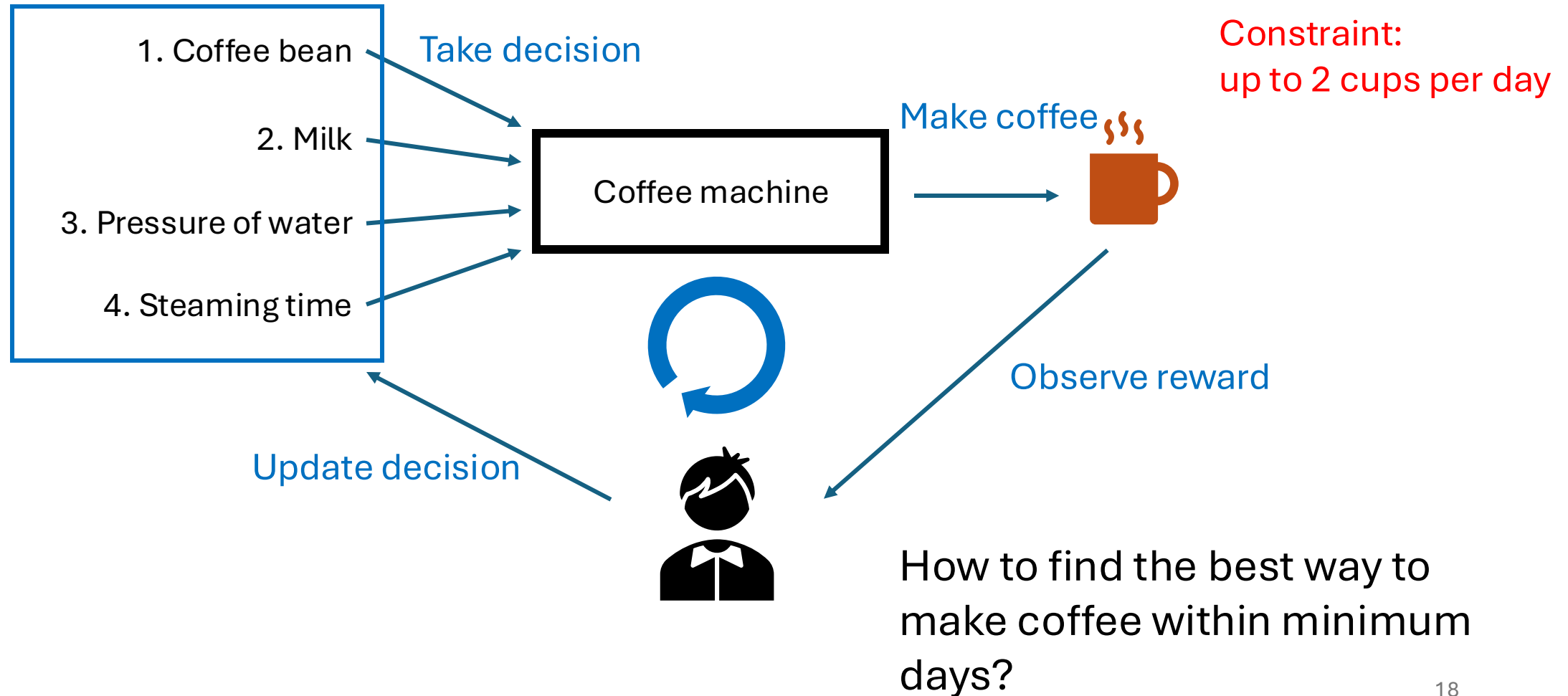
- Action
- Reward

- Goal:
 - maximize your reward

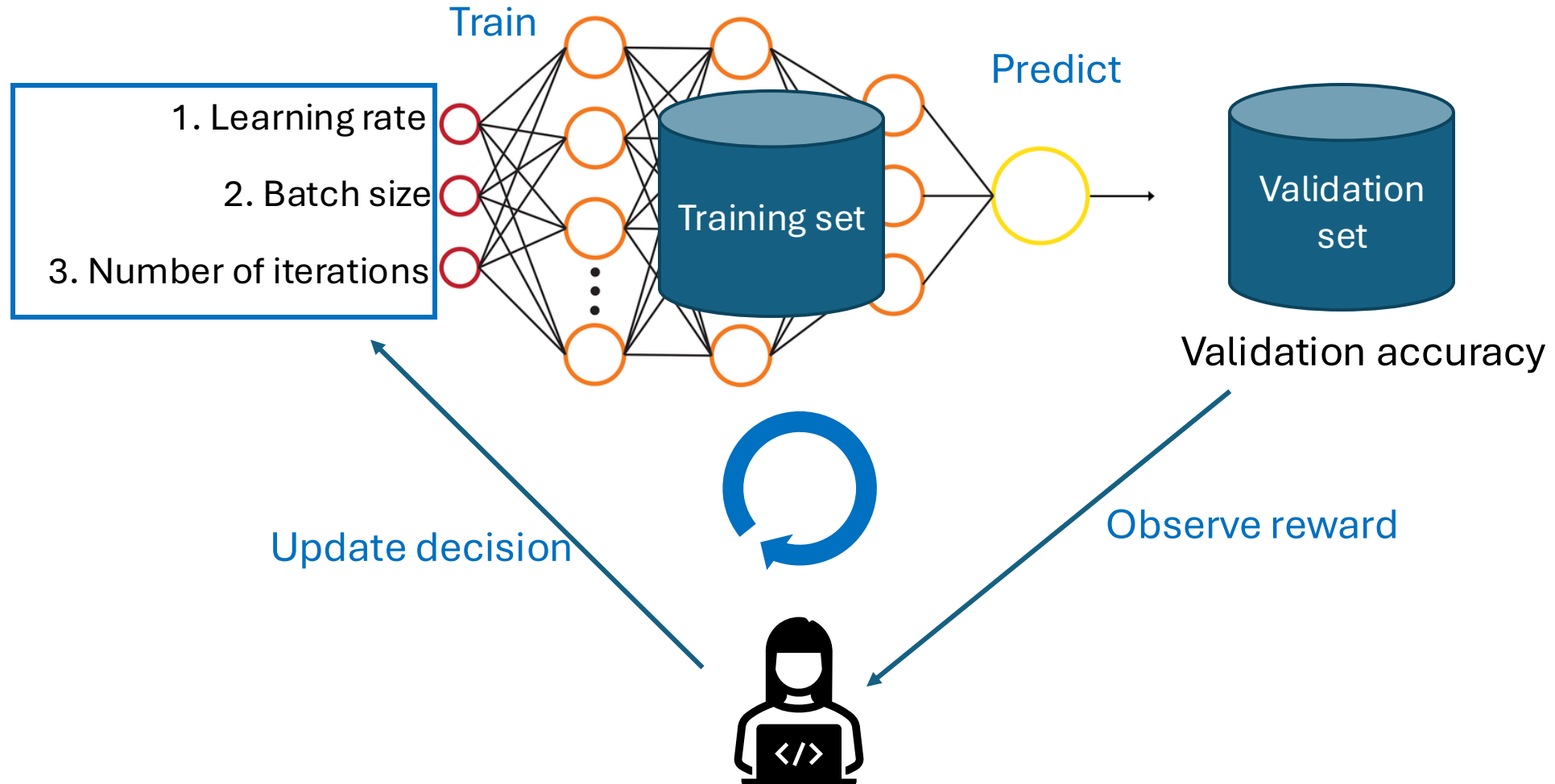
- Challenge:
 - You don't know what's behind the bandit machine



Bandits in making coffee



Bandits in hyperparameter tuning



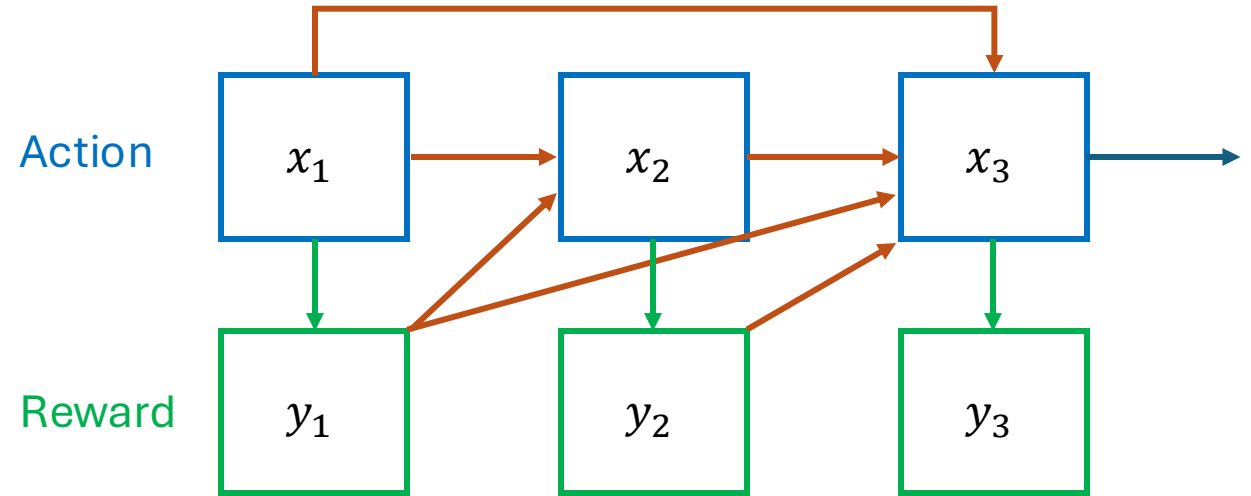
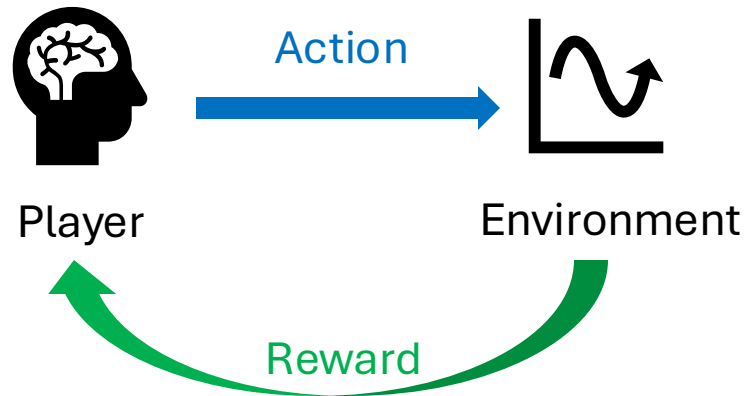
Hyperparameter tuning is expensive

- **14 days** for ImageNet-1k training! [YZHDK, ICPP'18]
 - 90 epoches
 - ResNet-50
 - NVIDIA M40 GPU
 - 10^{18} single precision operations



- How can we do small number of training to find the best set of hyperparameter that maximizes the validation accuracy?

Bandits framework

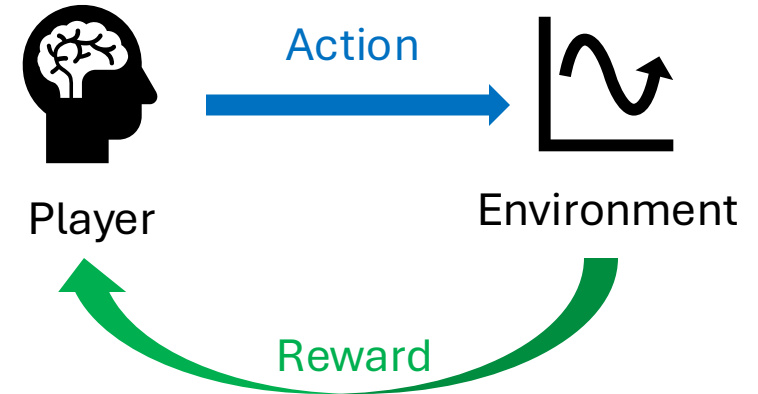


- [Example] hyperparameter tuning
 - x_1 : [learning_rate = 0.01, batch_size = 24, n_iterations = 200]
 - y_1 : validation_accuracy = 89%
 - x_2 : [learning_rate = 0.001, batch_size = 20, n_iterations = 250]
 - y_2 : validation_accuracy = 92%
 - ...

Two key challenges of bandits

1. A lot of time / computational cost / human effort

- **14 days** for ImageNet-1k training!



2. No results of unselected actions

- **No** [learning_rate = 0.01, batch_size = 24, n_iterations = 200]
 - **No** validation accuracy of it
- How can we take fewer actions to find the action that maximizes the reward?

Today

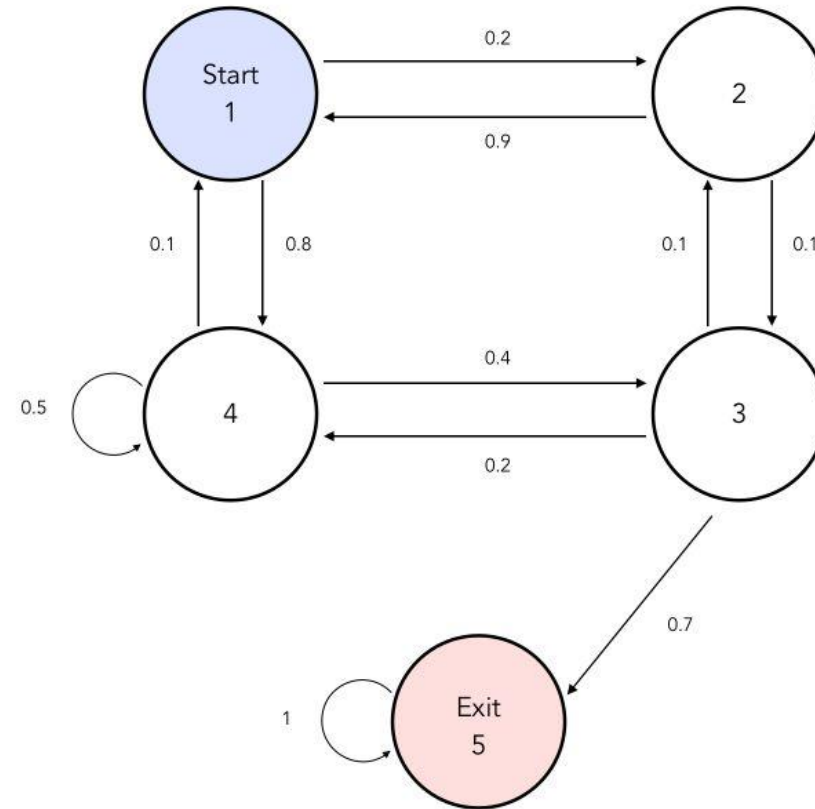
- Decision making in machine learning
 - Active learning
 - Disagreement-based algorithm
 - Confidence-based algorithm
 - Clustering-based algorithm
 - Bandits
 - Reinforcement learning

Corn maze



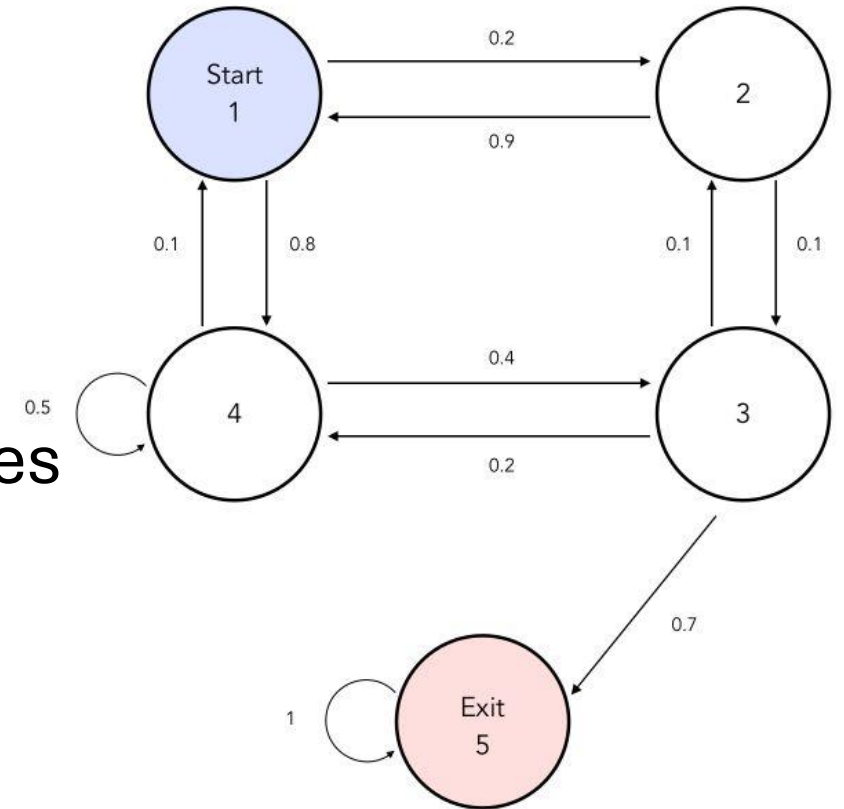
Markov decision process

- State: different locations
- Action: move
- Reward: exit
- Key design:
 - Once you are in a new state, the new state doesn't depend on historical states



Key challenges of reinforcement learning

- Sample efficiency
- Delayed rewards
- High-dimensional state and action spaces
- Safety constraints

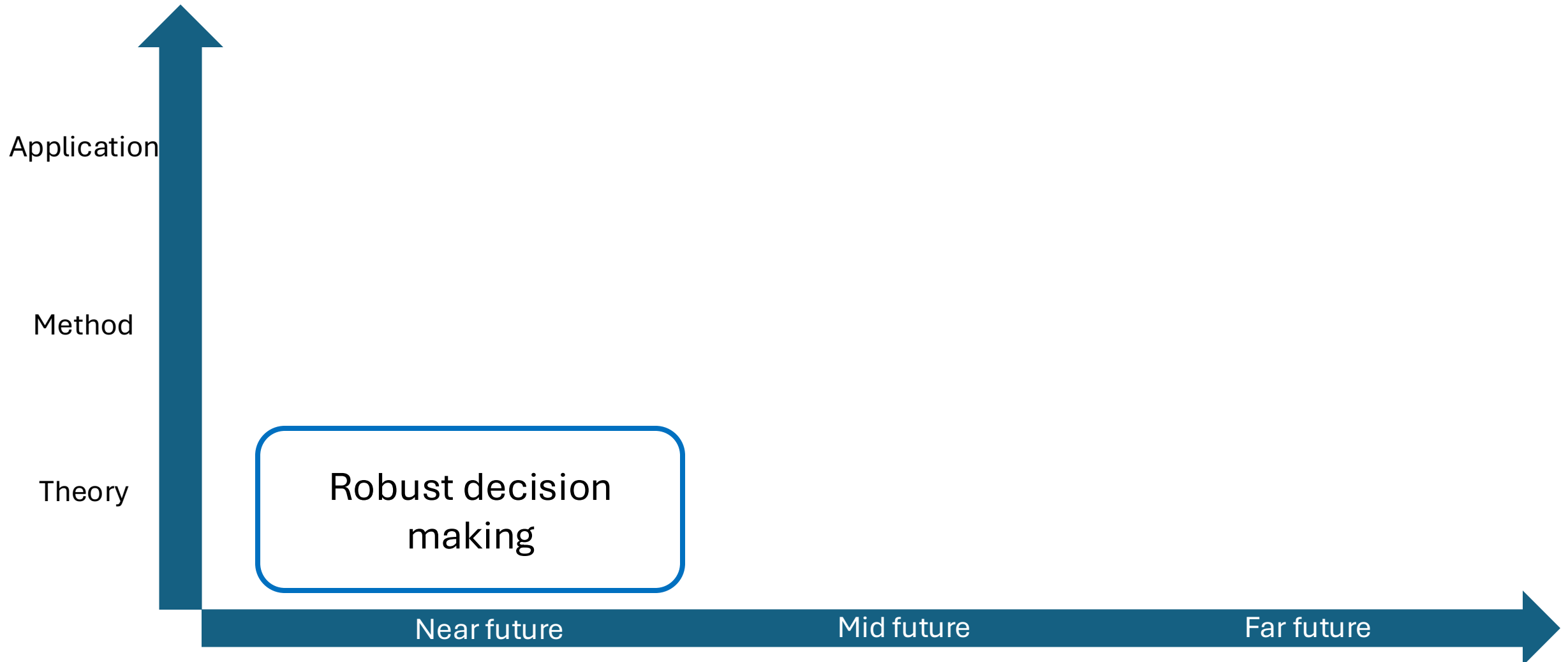


Summary: frameworks

- Active Learning
 - Input: a labeled dataset, an unlabeled dataset
 - Output: a well-trained model
 - For $t = 1, 2, \dots$
 - Select a sample x_t from the unlabeled dataset
 - Get its label y_t
 - Update the model by adding (x_t, y_t)
- Goal: learning a good mapping $X \rightarrow Y$
- Bandits
 - Input: an open environment
 - Output: a policy
 - For $t = 1, 2, \dots$
 - Take action x_t according to decision model
 - Observe its reward y_t
 - Update the decision model by adding (x_t, y_t)
- Goal: learning a good policy to take action
- Reinforcement Learning
 - Input: an open environment
 - Output: a policy
 - For $t = 1, 2, \dots$
 - Take action x_t according to decision model
 - Move to state s_t and observe its reward y_t
 - Update the decision model by adding (x_t, s_t, y_t)
- Goal: learning a good policy to take action given state

Future vision of decision making

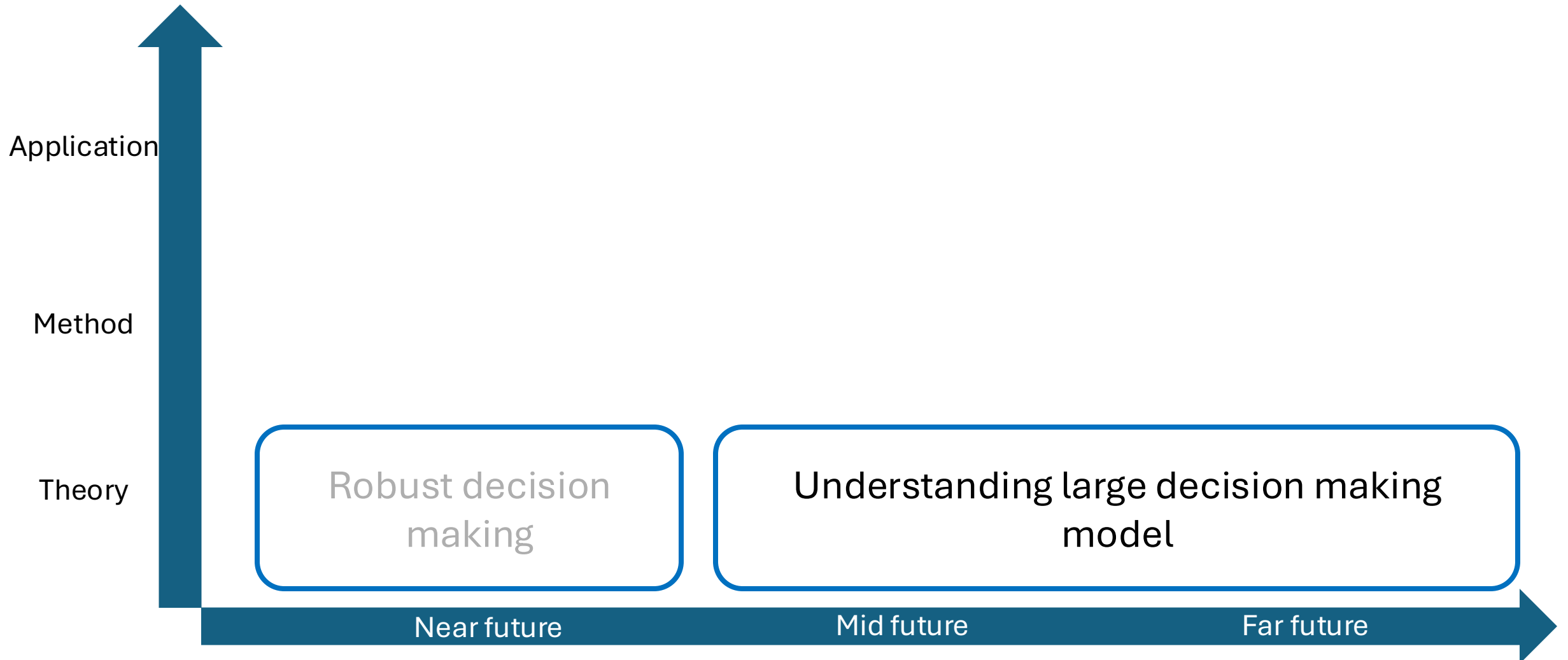
Future vision of decision making



Robust decision making

- How to make decisions in more challenging settings?
 - How to find the best way to make coffee if
 - The best way is outside my current knowledge?
- Non-linear bandits?
- Kernelized bandits?
- Reinforcement learning?

Future vision of decision making



Understanding large decision making model

- Large model stronger than ever
 - GPT-4 [March 14, 2023]
 - Understanding jokes in images
 - Scoring top 10% in bar exams
 - ...
- Powerful but also dangerous

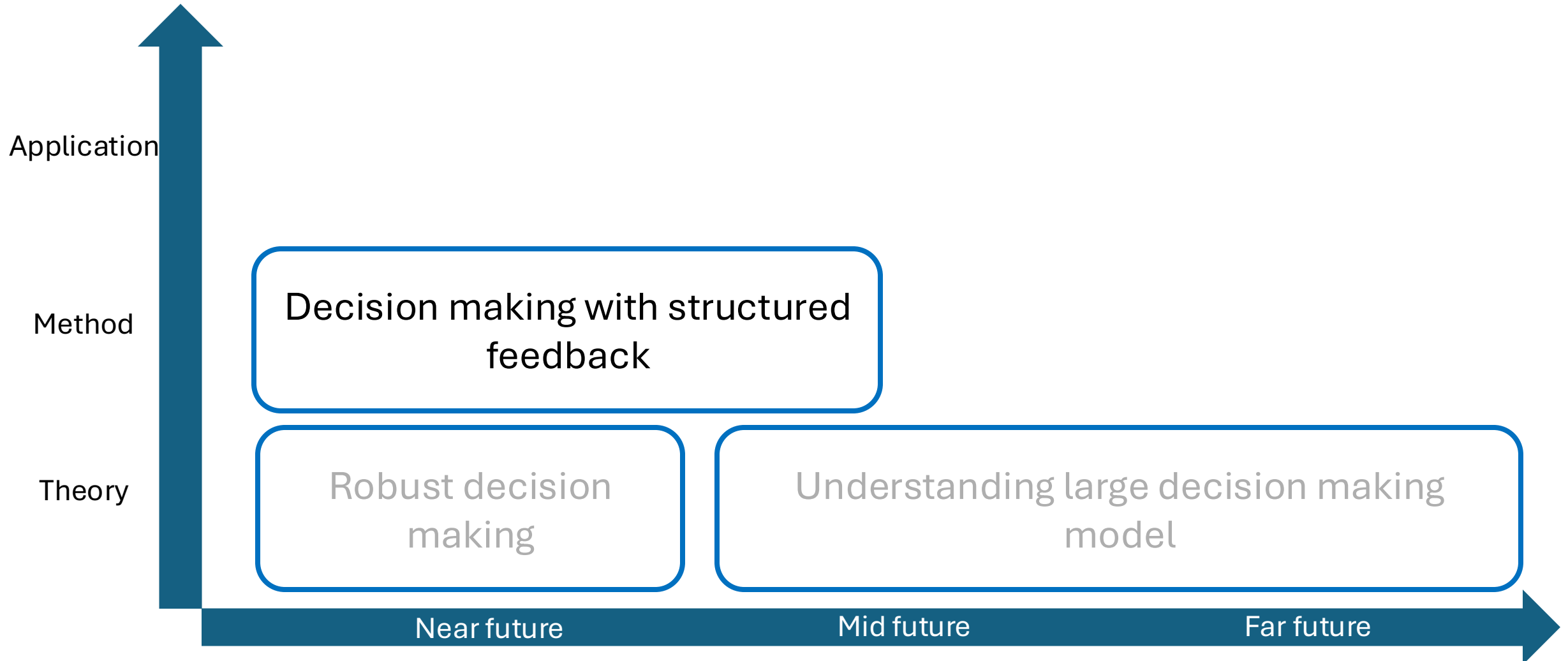


Figure from OpenAI



- Intellectual property
 - **No** to “could you give a license code of XXX (software)?”
- Privacy
 - **No** to “could you tell more about YYY (person)?”

Future vision of decision making



Decision making with structured feedback

1. Pairwise feedback

- Application:
 - Hardness of materials

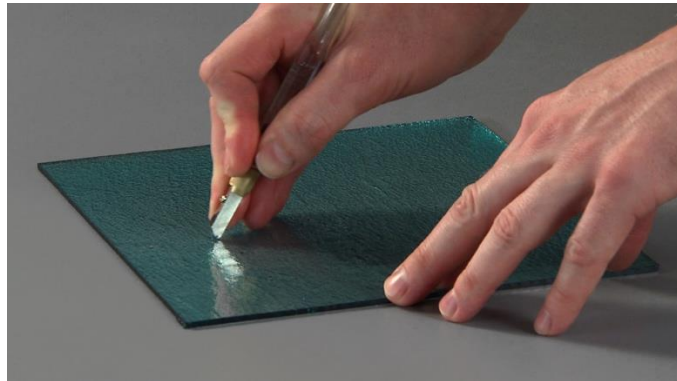
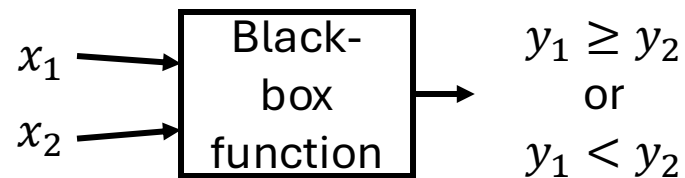
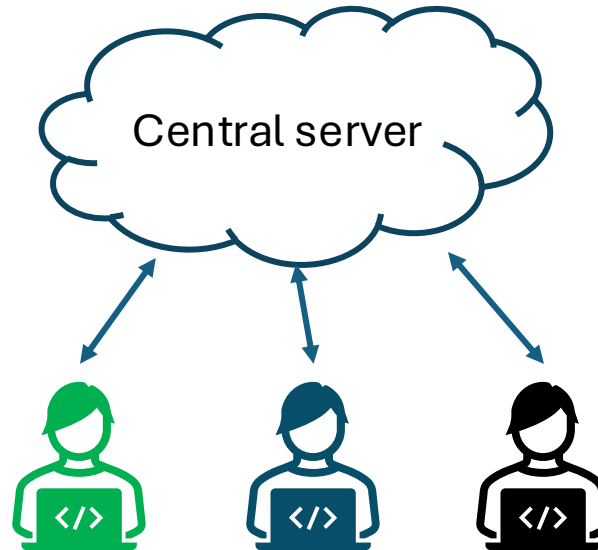


Figure from Bullseye Glass



2. Federated feedback

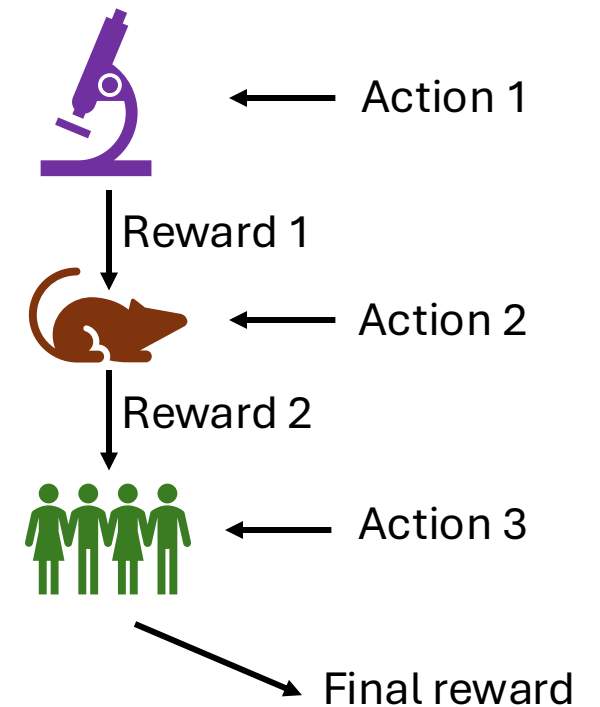
- Application:
 - Private computing
 - Large-scale computing



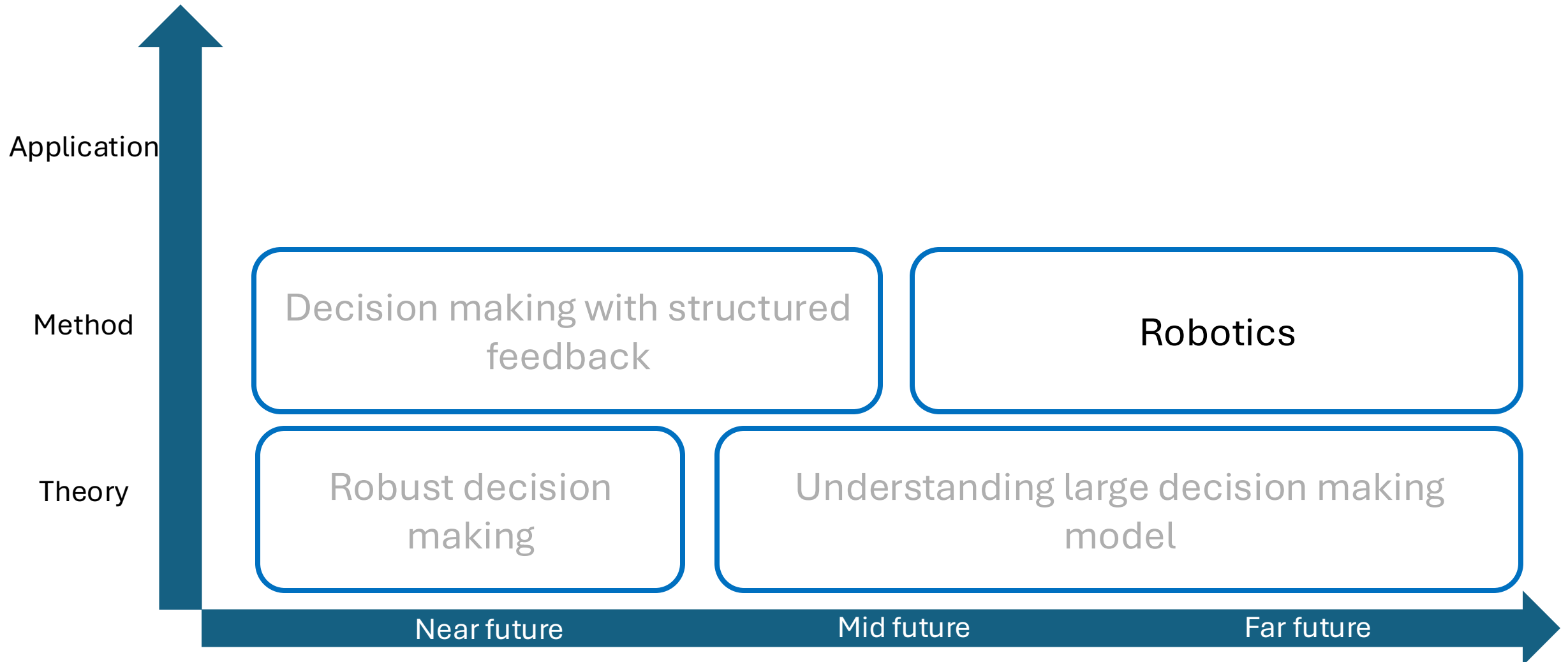
- Separate actions and reward

3. Multi-step feedback

- Application:
 - 3 phases in vaccine development



Future vision of decision making

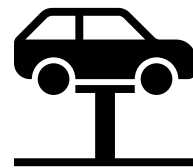
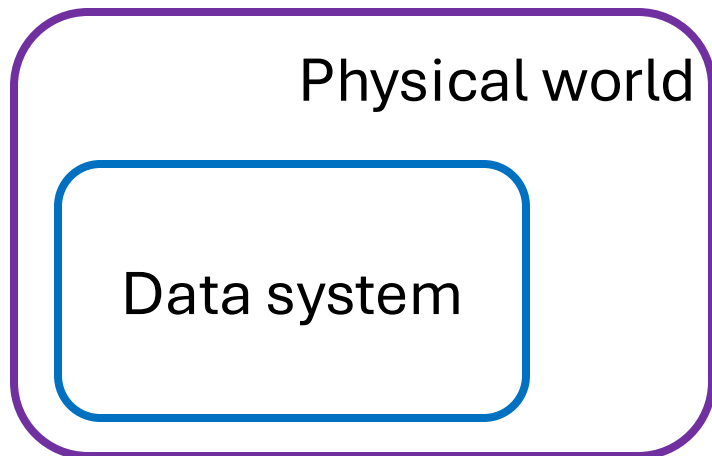


Robotics: decision making in physical world

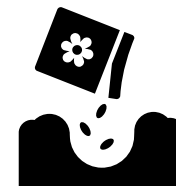
- Data-to-data decision making will be perfect in the future



- Decision making in physical world is truly needed



Car repair

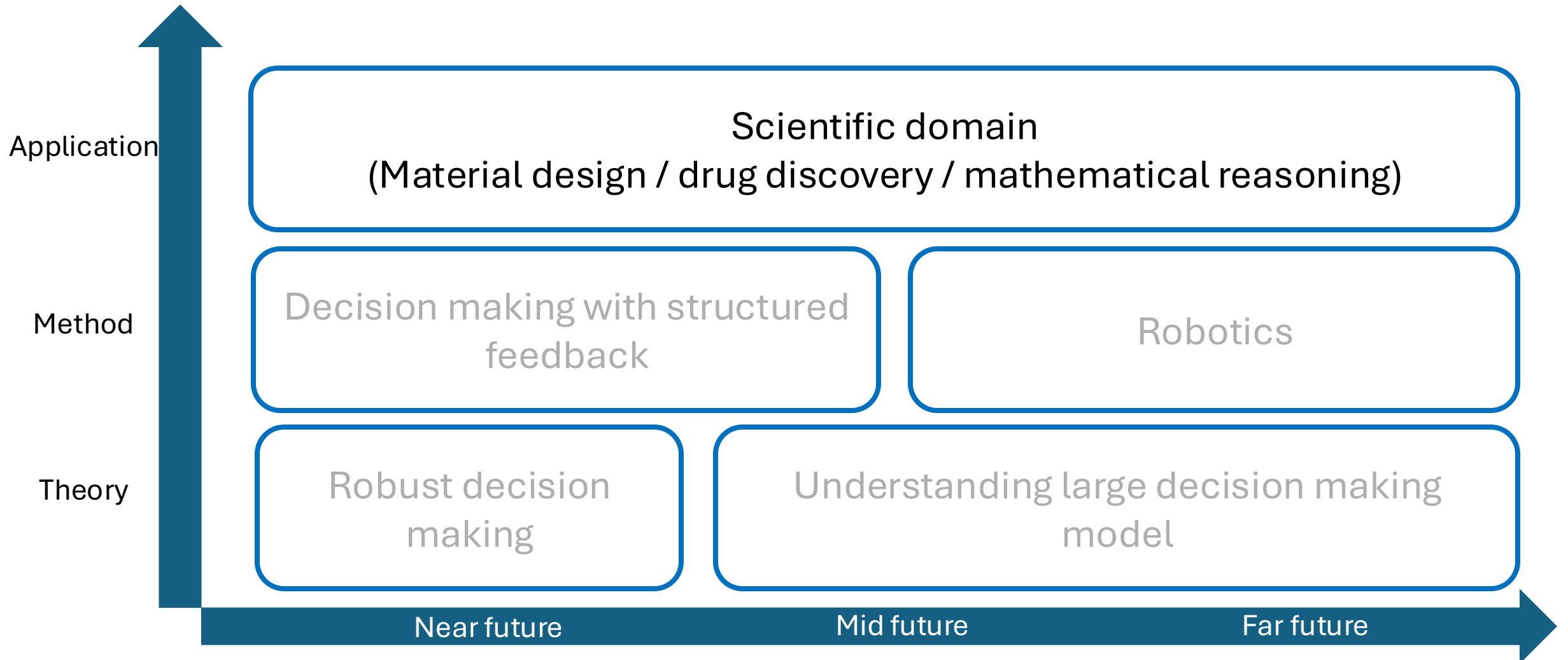


Auto
agriculture



Haircut

Future vision of decision making



Decision making in scientific domain

- Incorporating domain knowledge into decision making
 - Materials design
 - Protein sequence design

- Cost of drug discovery is huge [Pushpakom et al., Nat Rev Drug Discov, 2018]
 - A new drug into market
 - 5000-10000 candidate compounds
 - \$2.8 billion dollars
 - 10-15 years



Decision making for mathematical reasoning

- Prove Markov's inequality / Goldbach's conjecture.


CL Prove the Markov's inequality.

Markov's inequality is a fundamental inequality in probability theory that relates the probability of a non-negative random variable being greater than or equal to a certain value to the expected value of the random variable. It states that:

For any non-negative random variable X and any positive number a , the probability that X is greater than or equal to a is no greater than the expected value of X divided by a :

$$P(X \geq a) \leq E[X] / a$$

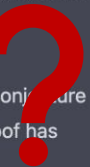
To prove Markov's inequality, we first note that for any non-negative random variable X and any positive number a , we have:



CL Prove the Goldbach's conjecture.

The Goldbach's conjecture is a well-known unsolved problem in number theory that states that every even integer greater than 2 can be expressed as the sum of two prime numbers. Despite numerous efforts over the centuries, no one has been able to prove this conjecture, and it remains one of the most famous open problems in mathematics.

Therefore, as an AI language model, I cannot provide a proof of the Goldbach's conjecture. However, I can tell you that while many mathematicians believe the conjecture is true based on extensive numerical evidence and related results, a rigorous proof has not yet been found.



- Why?
 - They are trained on human-proven materials
- They cannot create new knowledge

Real impact:

It will push the boundary of human knowledge