Optimization algorithms for Artificial Intelligence: Introduction

Alberto TONDA, Senior Researcher (DR), <alberto.tonda@inrae.fr>

UMR 518 MIA-PS (Applied Mathematics and Computer Science)
INRAE, AgroParisTech, Université Paris-Saclay
Institut des Systèmes Complexes, Paris-Ile-de-France

Outline

- What is this class about?
- Who am I?
- What is optimization?
- What is Artificial Intelligence?
- Is optimization a kind of Artificial Intelligence?
- What is the relationship between AI and optimization?
- Why are we still here? *Just to suffer? Every night, I can feel my leg... And my arm... even my fingers... The body I've lost... the comrades I've lost... won't stop hurting... It's like they're all still there. You feel it, too, don't you? I'm gonna make them give back our past!*

What is this class about?

- Optimization!
  - Overview of optimization techniques, when/how to use them
  - How these techniques power modern Artificial Intelligence
  - Optimization to improve performance of AI methods
- At the end of the class, you should know
  - What techniques are more appriopriate for different problems
  - How (several) AI systems work, especially for Machine Learning
  - Hyperparameter optimization for your AI applications
  - Keywords for further research (e.g. Neuro-symbolic, AutoML)

What is this class about?

- Optimization!
  - **Overview of optimization techniques, when/how to use them**
  - How these techniques power modern Artificial Intelligence
  - **Optimization to improve performance of AI methods**
- At the end of the class, you should know
  - What techniques are more appriopriate for different problems
  - How (several) AI systems work, especially for Machine Learning
  - Hyperparameter optimization for your AI applications
  - Keywords for further research (e.g. Neuro-symbolic, AutoML)

(Tentative) Schedule

April 1, Amphi IV
- Optimization: introduction - Continuous optimization
Exercises
- Linear programming

April 3, Amphi VI
- Discrete optimization
Exercises
- Multi-objective optimization
Exercises
- Optimization of structures
Exercises

April 8, Amphi IV
- Optimization in ML
Exercises
- Hyperparameter optimization

April 8, Amphi IV
Discussion
- Recent developments

(Tentative) Schedule

April 1, Amphi IV
- Optimization: introduction - Continuous optimization
Exercises
- Linear programming

April 3, Amphi VI
- Discrete optimization
Exercises
- Multi-objective optimization
Exercises
- Optimization of structures
Exercises

April 8, Amphi IV
- Optimization in ML
Exercises
- Hyperparameter optimization

April 8, Amphi IV
Discussion
- Recent developments

Prepare a few slides to **present your problem**, then we discuss how to frame it as an **optimization problem**

What happened in past iterations

April 1, Amphi IV
- Optimization: introduction - Continuous optimization
Exercises
- Linear programming

April 3, Amphi VI
- Discrete optimization
Exercises
- Multi-objective optimization
Exercises
- Optimization of structures
Exercises

April 8, Amphi IV
- Optimization in ML
Exercises
- Hyperparameter optimization

April 8, Amphi IV
Discussion
- Recent developments

Students asked me questions after class on during pauses (shyness)? I asked for feedback

What happened in the past

What is this class about?

- **General ideas** are relatively easy to grasp
- **Details** are complicated and require longer study
- Get the general idea, try to understand if it fits your problem

Who am I?

- Career
  - Bachelor and Master in Computer Science Engineering
  - Ph.D. from Politecnico di Torino, Italy, in 2011
  - Permanent researcher in France since late 2012 (INRAE)
  - Senior researcher since 2023
- Research interests
  - Stochastic optimization, ML (xAI)
  - Applied to biological/agri-food data
  - Research: applied + algorithms

What is optimization?

- Nearly every choice is an optimization problem
  - Shape of a car to minimize wind resistance
  - Values of the weights of a neural network to best perform a task
  - Weight distribution in a plane to minimize shaking
  - Pick a stock market portfolio to maximize revenue
  - Trace route inside a city to reach a point as fast as possible
  - Choose career that makes you satisfied and happy
  - ...

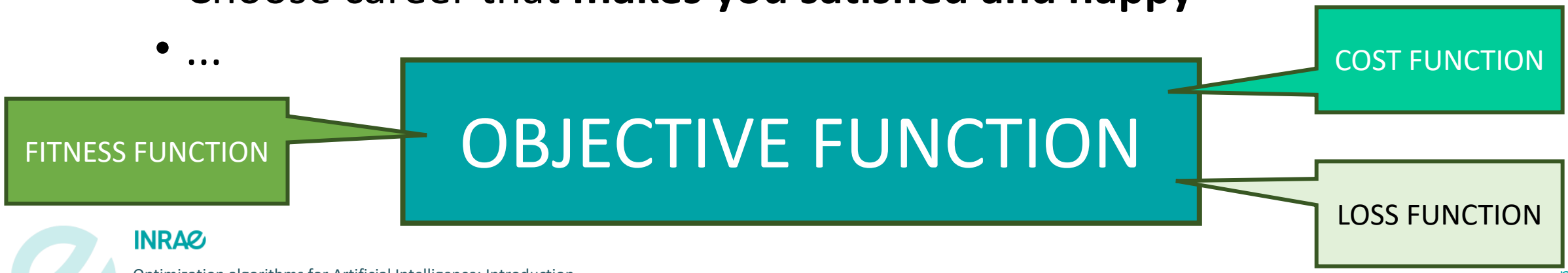
What is optimization?

- Nearly every choice is an optimization problem
  - Shape of a car to **minimize wind resistance**
  - Values of the weights of a neural network to **best perform a task**
  - Weight distribution in a plane to **minimize shaking**
  - Pick a stock market portfolio to **maximize revenue**
  - Trace route inside a city to **reach a point as fast as possible**
  - Choose career that **makes you satisfied and happy**
  - …

OBJECTIVE FUNCTION

What is optimization?

- Nearly every choice is an optimization problem
  - Shape of a car to **minimize wind resistance**
  - Values of the weights of a neural network to **best perform a task**
  - Weight distribution in a plane to **minimize shaking**
  - Pick a stock market portfolio to **maximize revenue**
  - Trace route inside a city to **reach a point as fast as possible**
  - Choose career that **makes you satisfied and happy**
  - …



What is optimization?

- Objective function
  - Measure of goodness of a candidate solution
  - Quantitative, not qualitative (unless we can somehow sort it)
  - Good candidate solutions are usually close to other good solutions
  - If you pick the wrong objective function, you are screwed
- Candidate solutions
  - Possible inputs of the objective function
  - High-level representation that includes all possible solutions
  - Example: $\{x_0, x_1, \ldots, x_n\}$ with $x_i \in [0,1] \cap \mathbb{R}$

What is optimization?

- Given an objective function $y = f(\boldsymbol{x})$, find $\boldsymbol{x}^*$ such that

$$\boldsymbol{x}^* = \underset{\boldsymbol{x}}{argmax}(y) \qquad \boldsymbol{x}^* = \underset{\boldsymbol{x}}{argmin}(y)$$

- Find the values of $\boldsymbol{x}$ that **maximize** or **minimize** y

What is optimization?

What is an algorithm?

What is an algorithm?

- Series of steps to achieve an objective
  - Bake a cake
  - Sort the elements inside an array, by increasing value
  - Find the best possible input value for a given function
- Algorithmic complexity
  - Time to complete objective often depends on **size of input**
  - Baking ten cakes takes more time than baking one
  - Sorting a large array takes more time than a small one

What is an algorithm?

- Algorithmic complexity is important for **expectations**
  - In particular, worst-case scenario (upper bound)
  - Notation: $O(f(n))$ where n is the size of the input

- Polynomial time vs Super-polynomial time
  - $O(n^k)$ vs $O(2^{n^k})$ or $O(n!)$, “easy” vs “difficult”
  - In practice, polynomial time can be used for large instances
  - Super-polynomial time can only be used for small input size

What is an algorithm?

- Computer scientists often look at worst-case scenario
  - In some cases, average-case scenario might be better
  - E.g. worst-case scenario exponential time, average polynomial

- Most real-world problems are Nondeterministic Polynomial
  - Checking if a solution is correct in polynomial time
  - Finding the best solution takes super-polynomial time

What is Artificial Intelligence?

What is Artificial Intelligence?

- John McCarthy, one of the founding fathers of Artificial Intelligence
- *«I invented this term Artificial Intelligence [...] because [...] we were **trying to get money** for a summer study [...] in 1956 [...]»*

What is Artificial Intelligence?

- Short answer, there is no clear definition
  - We do not have a good definition of *intelligence*, so...
  - Broadly speaking, AI defines a *field* more than a *method*
  - Machine learning, reinforcement learning, symbolic AI, ...
- Tentative definitions (there is no agreement)
  - «When a non-biological being successfully completes a task commonly believed to require biological intelligence»
  - «Perceiving, synthesizing, and inferring information»
- How do we *measure* intelligence?

What is Artificial Intelligence?

NARROW / WEAK

Focused on a specific task

- Symbolic AI
  - E.g. rule-based systems
- Machine learning
  - Supervised, unsupervised
  - Natural language processing
  - Image recognition/segmentation
- Reinforcement learning
- Neuro-symbolic AI

GENERAL (AGI)

Can perform any type of (human?) task

- Does not exist (...yet?)
- Closest thing is NLP: Large Language Models (LLM) like ChatGPT

Symbolic AI

- Symbolic manipulation
  - Reality is *continuous* (with good approximation)
  - Symbols are *discrete*, and humans are good at using them

Symbolic AI

- Symbols seem normal and natural, map into the real world (in linguistics, it's called *extension*)
- Natural language is a powerful human symbol manipulator
- However, there is chaos hidden under the surface
  - What is the reality of a *river*?
  - What is the reality of a *chair*?
  - What is the reality of a *number*?

Symbolic AI

- Symbol can be hard to define, but we grasp it intuitively
  - It's an old, old problem: see Plato and Diogenes
  - *Entire fields of research* on this (neuroscience, cognitive sciences, neurolinguistics, …)
- "Explaining" symbols to AI is harder yet
- Issues with "common sense"
- Reached limits in the 1980s

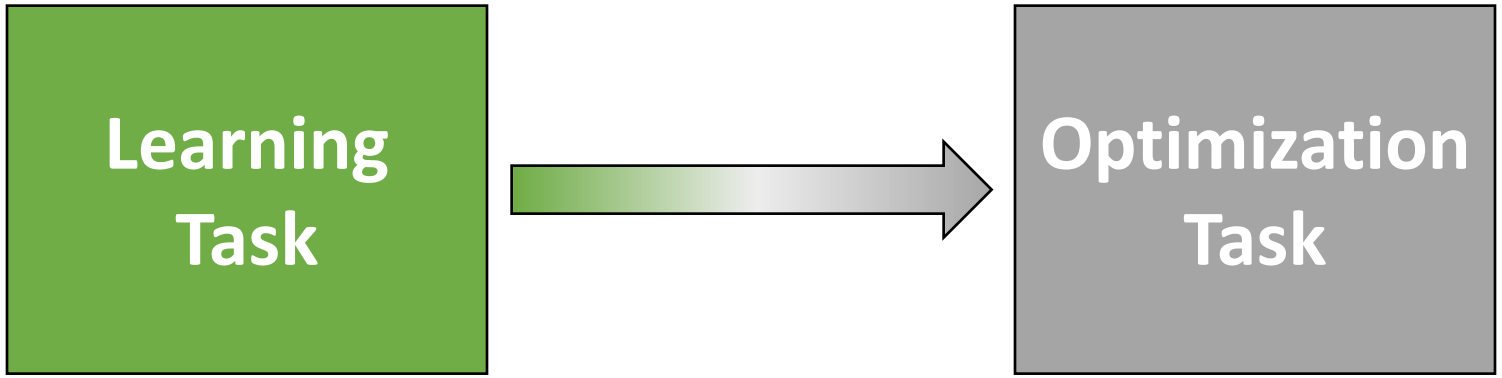
Symbolic AI

- In practice, find or exploit human-readable rules
  - Expert systems (“if-then-else” rules)
  - Knowledge graphs, linking entities with relationships
  - First-order logic rules
  - Decision trees (that are also considered part of ML!)

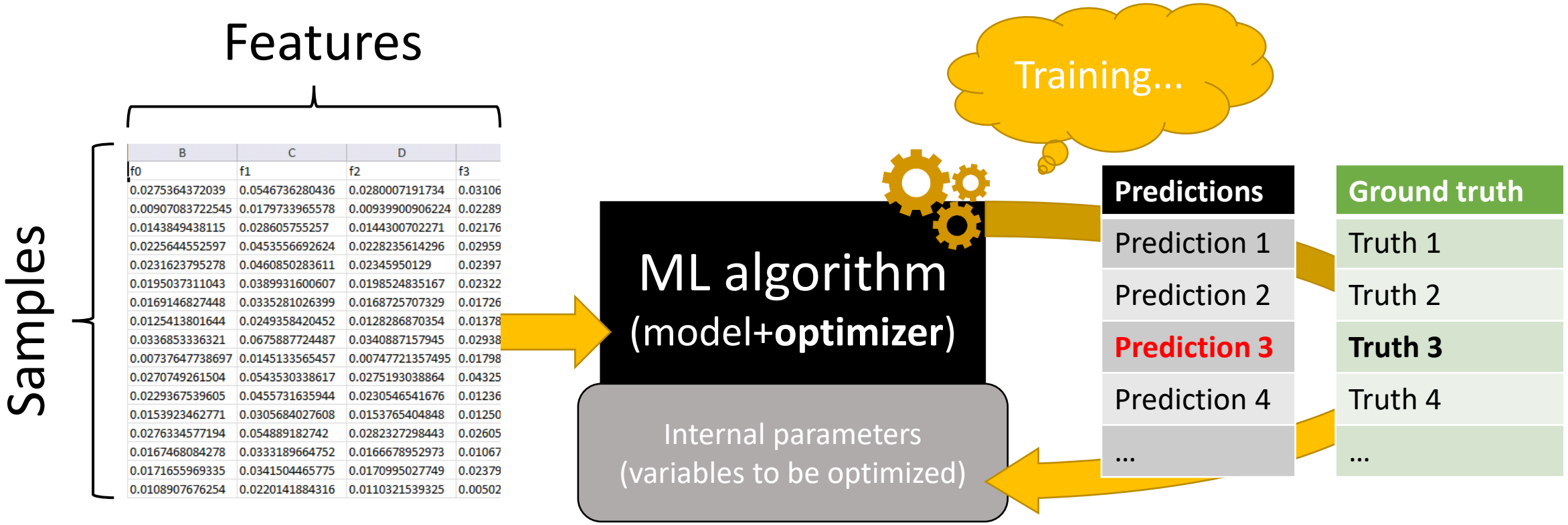
- Before the advent of ML, considerable success stories
- Symbolic AI is still in use, paired with ML

Machine learning

- Learn a task directly from examples
  - No need for symbols, just large quantities of data
  - *Samples* (rows) and *features* (columns)
- "Dirty secret" of ML: it's mostly optimization
  - Restate **learning task** as **optimization task**
  - Solve it relying on available (training) data



Machine learning (supervised)



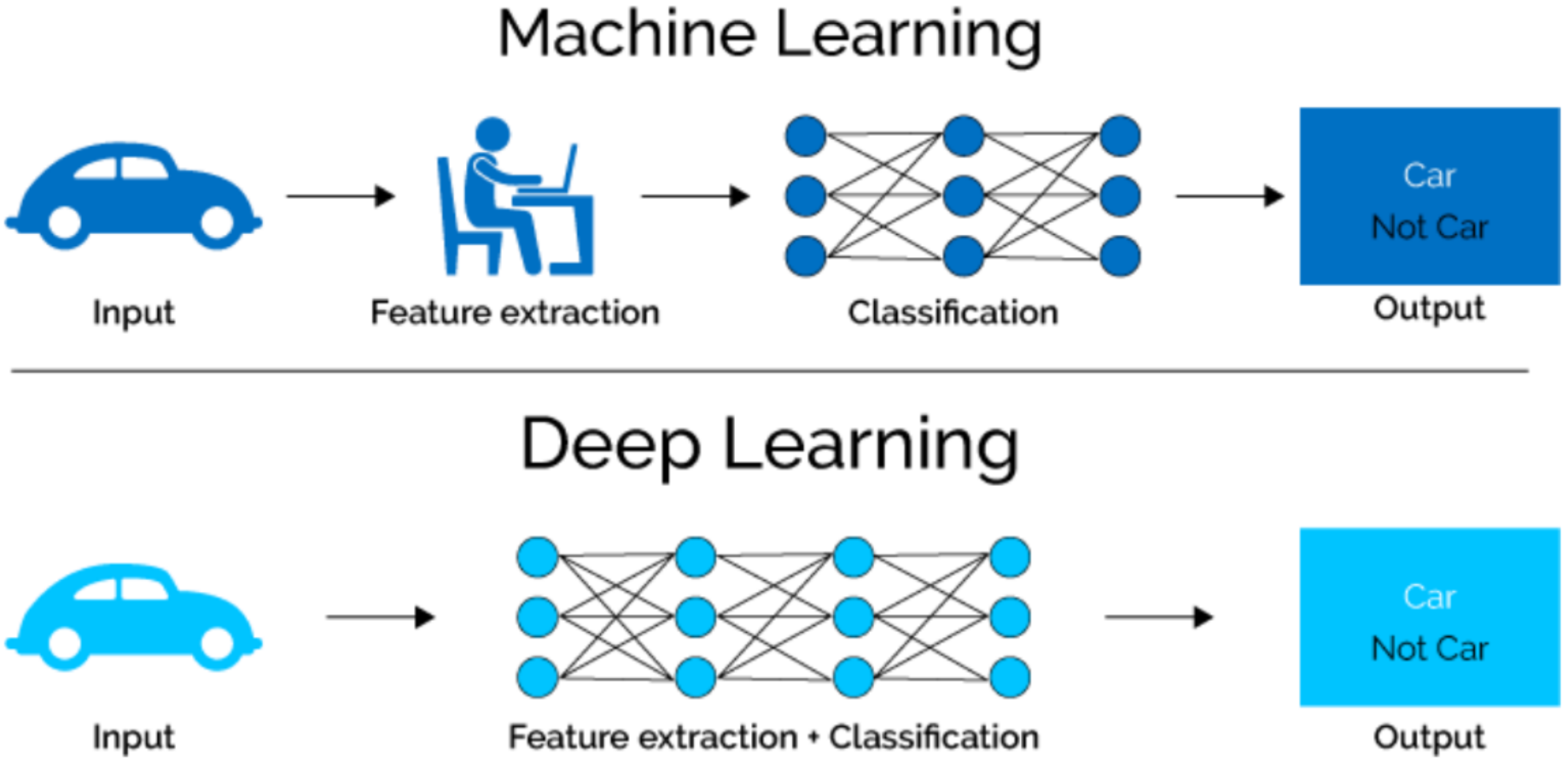
Machine learning

- Dominant paradigm since the 90s
  - First, artificial neural networks (ANNs or NNs)
  - Then, statistical learning algorithms
  - Decision trees (and ensembles of) and polynomial models
- Feature engineering
  - Works (still) well for **tabular data** (e.g. Excel spreadsheet)
  - Huge issues with **relational data** (e.g. images, text, sound…)
  - Hand-crafted features (not very successful)

Machine learning

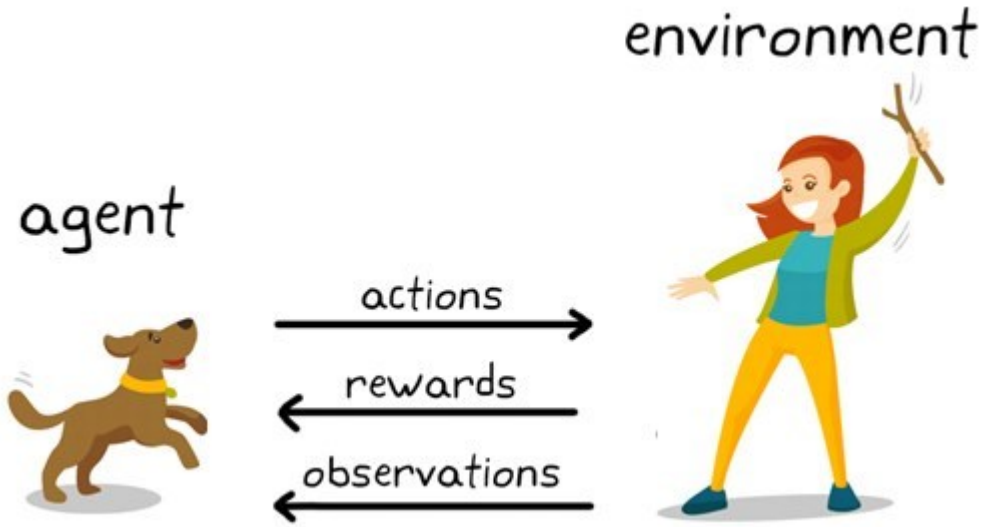
- Trade-off effectiveness/interpretability (black-box effect)
  - Good predictive models are extremely complicated
  - A single decision tree can be interpreted, 300 trees cannot
  - Same goes for polynomial models with 300 features
- Deep learning
  - General idea is that deep ANNs can automatically infer *features*
  - Convolutional NNs, Recurrent NNs, Transformers, …
  - Fantastic success stories for **relational data**!

Machine learning



Reinforcement learning

- Similar to ML, but not exactly
  - No value for a *single* decision; reward after *series* of decisions
  - Learn a **policy** which tells you what to do from the **state** you are in
  - Example: chess game; is trading a Queen for a Knight good? Well, it depends on the board **state**



Neuro-symbolic AI (NeSy)

- Might look complex, but the general idea is intuitive
  - Use neural/ML approach to map from data to symbols
  - Use symbolic AI to reason on symbols
  - (possibly) Go back to data using another neural/ML approach
  - Promise: **Effectiveness** (ML) + **Explainability** (Symbolic AI)
  - However, it's pretty hard to do, and problem-specific
- Interestingly, some of the biggest ML successes are NeSy
  - AlphaGo uses a mix of symbolic exploration and NN
  - AlphaFold is a mix of ~30 algorithms (some Symbolic, some ML)

Artificial General Intelligence

- Hypothetical artificial intelligent agent
  - "Can learn (rapidly and cheaply) to perform any task that a human or another animal could perform, with minimal amounts of errors"
  - **It does not exist**, there is no clear path towards it
  - Lots of people scared by apparently quick advances of AI
  - Even some real experts (!!!)
- **My opinion**
  - Existential risk is non-existent
  - Real risks are already here, from misuse/misunderstanding

Is optimization a kind of AI?

- Debatable, some experts would say “yes”
- My opinion: optimization is the **engine** of AI
- …this is almost sure for ML

Relationship between AI and optimization?

- Most “Intelligence” requires making <u>good choices</u>
  - Predict the next value in a time series, as precisely as possible
  - Correctly identify the human poses in a video
  - Make the best possible chess move, given the situation
  - Maximize your score in Super Mario
  - Generate a sequence of words that best follows the input
  - Create the painting that best corresponds to a written prompt

Relationship between AI and optimization?

- Most “Intelligence” requires making <u>good choices</u>
  - Predict the next value in a time series, **as precisely as possible**
  - **Correctly** identify the human poses in a video
  - Make the **best possible** chess move, given the situation
  - **Maximize** your score in Super Mario
  - Generate a sequence of words that **best follows the input**
  - Create the painting that **best corresponds** to a written prompt

Relationship between AI and optimization?

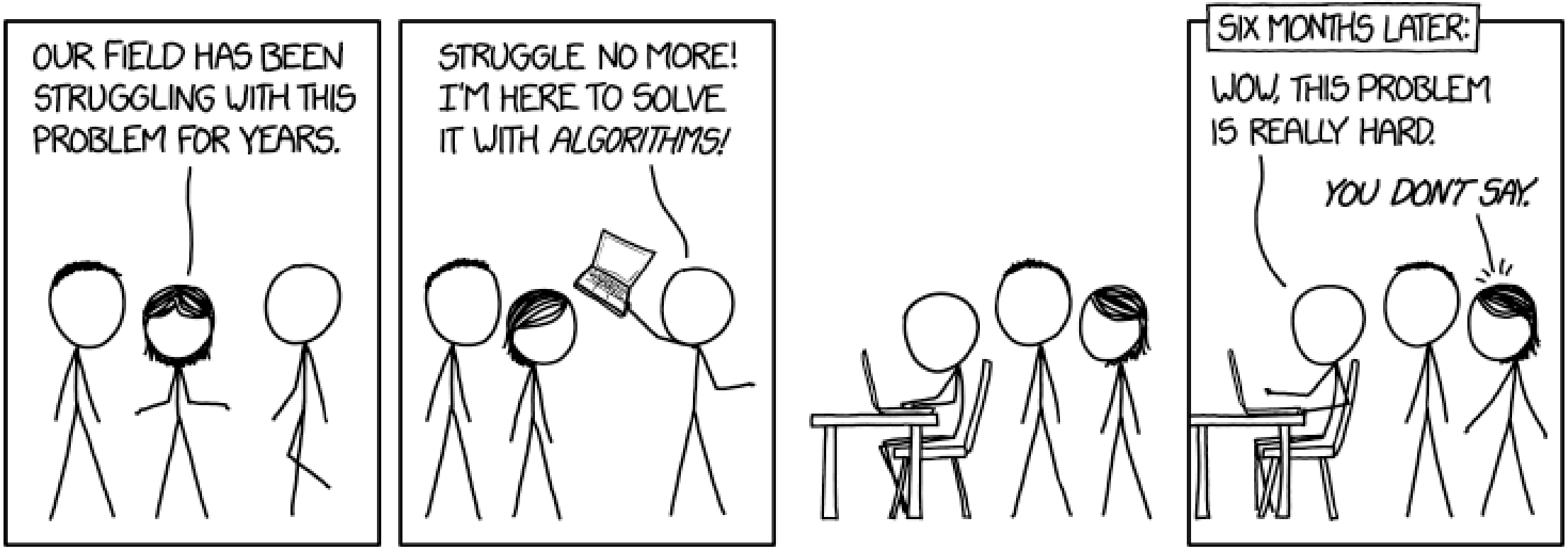
- AI: from many (many!!!) possible choices, pick the best
- If you can evaluate your choices, you can optimize
- (**AFAIK**) **All AI systems** include some type of optimization, with the possible exception of hand-crafted ones

Is it always good to optimize?

- Optimizing one objective might lead to undesired outcomes
  - Supply chain optimized for efficiency is *fragile*
  - “Robust optimization” considers perturbations, but how big?

- Multi-objective optimization can be helpful
  - Multiple conflicting objectives, improve one <-> worsen others
  - Find **optimal** trade-offs between objectives

Conclusions



"XKCD" by Randall Munroe www.xkcd.com

Questions?