CS 480

Introduction to Artificial Intelligence

April 16, 2024

Announcements / Reminders

- Please follow the Week 13 To Do List instructions (if you haven't already)
- Work on your last Written Assignment

- FINAL EXAM is on THURSDAY (04/25/2024)
  - IGNORE Registrar's FINAL EXAM date
  - Last week of classes! NOT finals week
  - Section 02: contact Mr. Charles Scott (scott@iit.edu) to make arrangements

Plan for Today

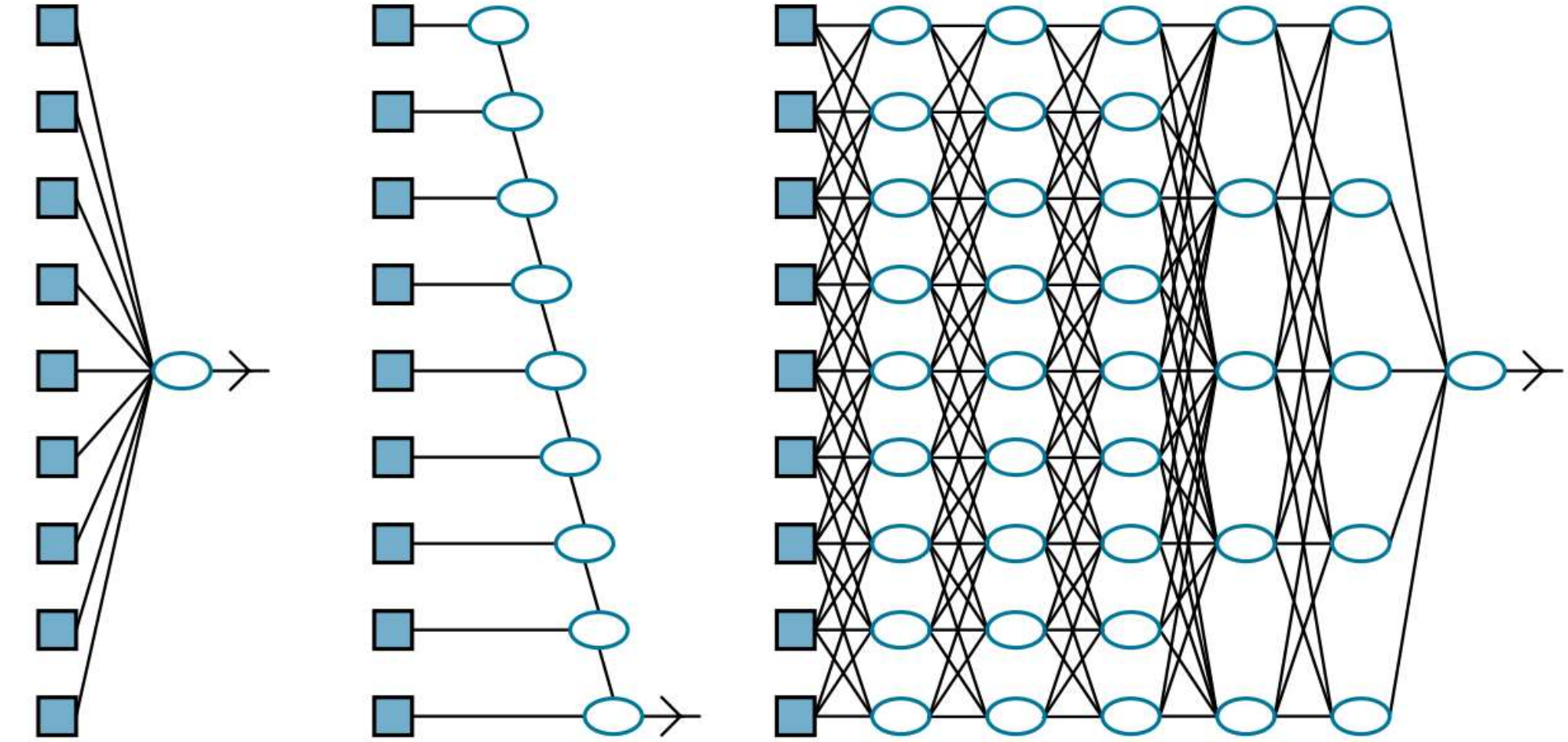
- A Casual Introduction to Machine Learning
- Deep Learning
- Bio-Inspired AI

Deep Learning

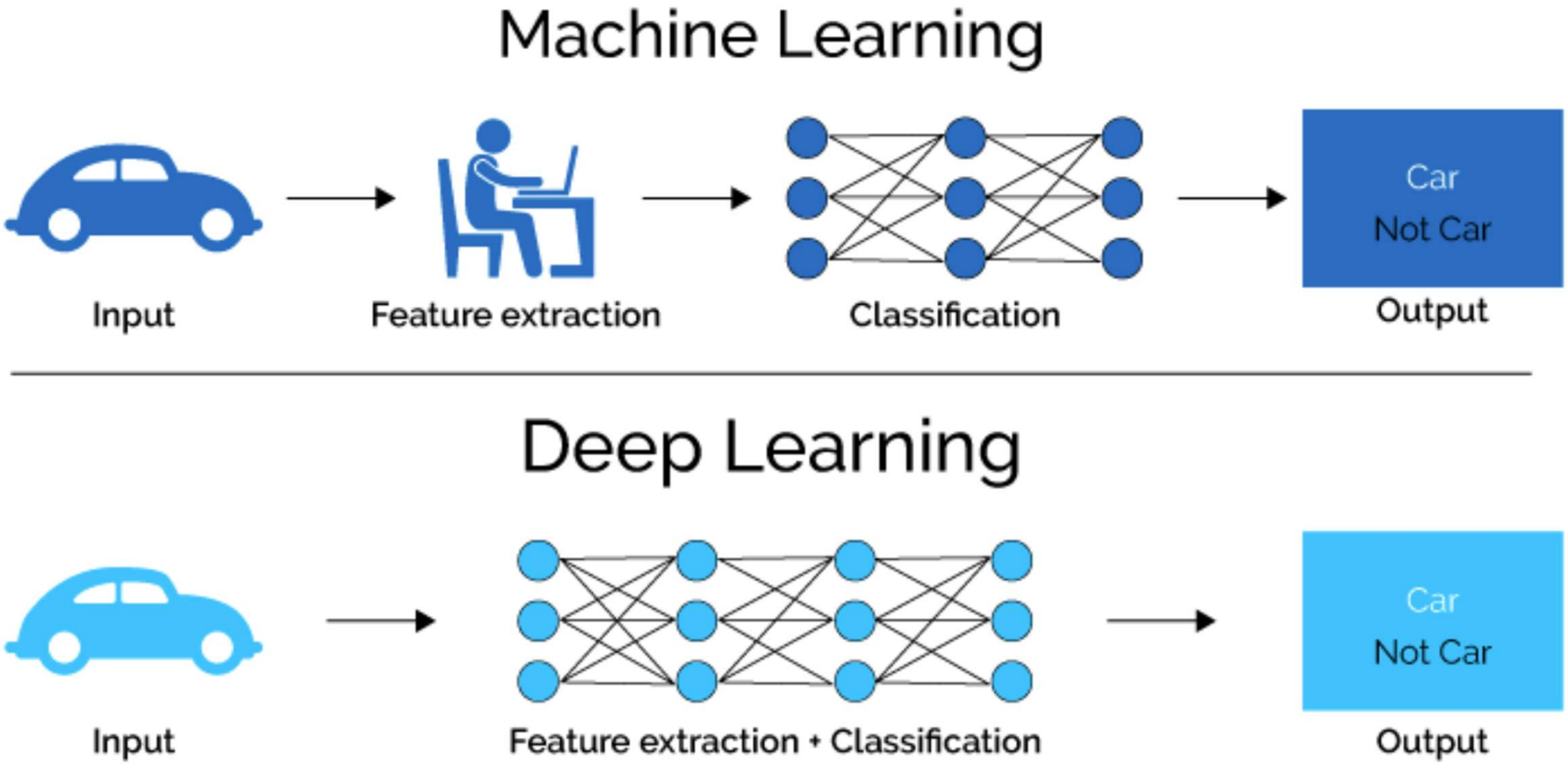
Deep Learning

Deep learning is a broad family of techniques for machine learning (also a sub-field of ML) in which hypotheses take the form of complex algebraic circuits with tunable connections. The word "deep" refers to the fact that the circuits are typically organized into many layers, which means that computation paths from inputs to outputs have many steps.

Shallow vs. Deep Models

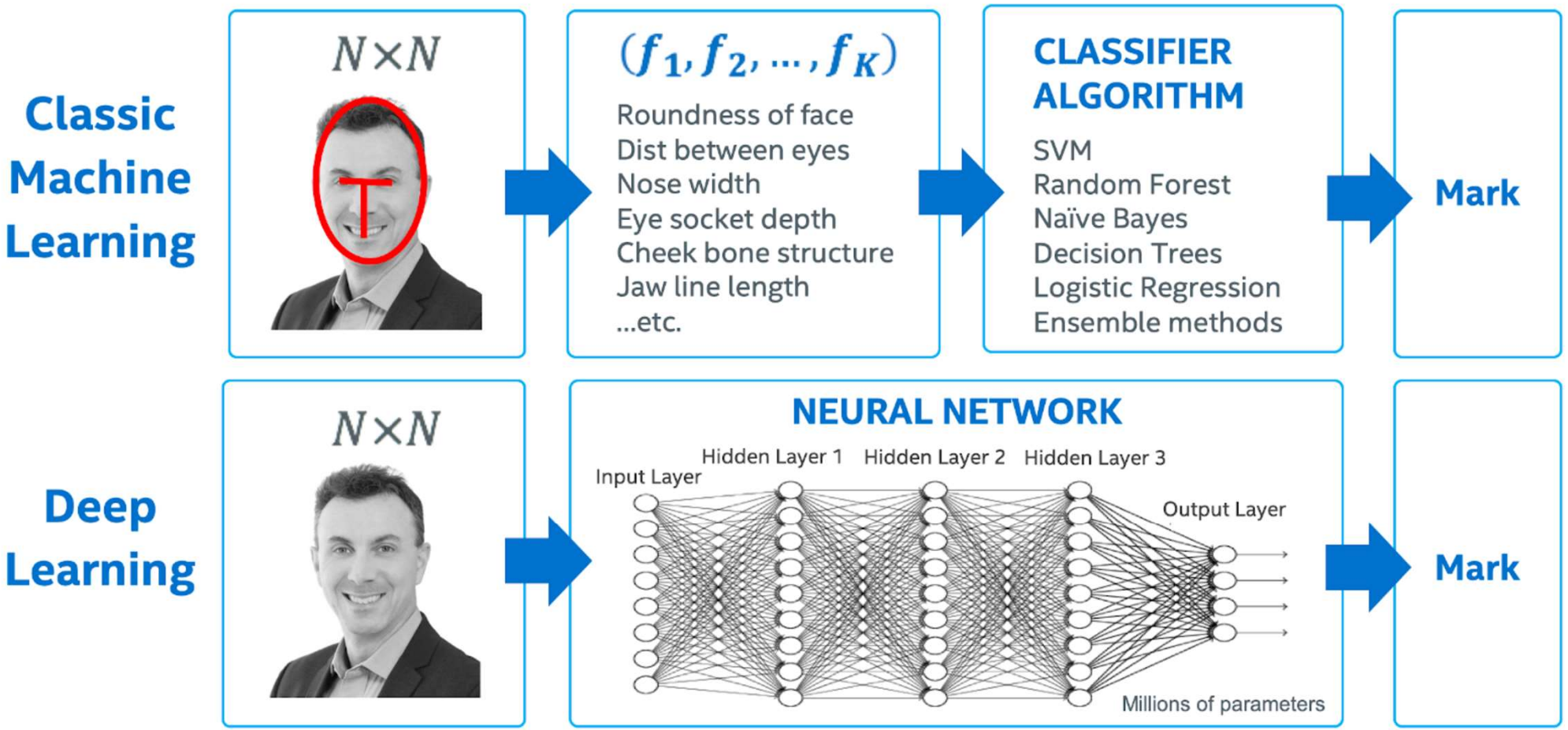


Machine Learning vs. Deep Learning



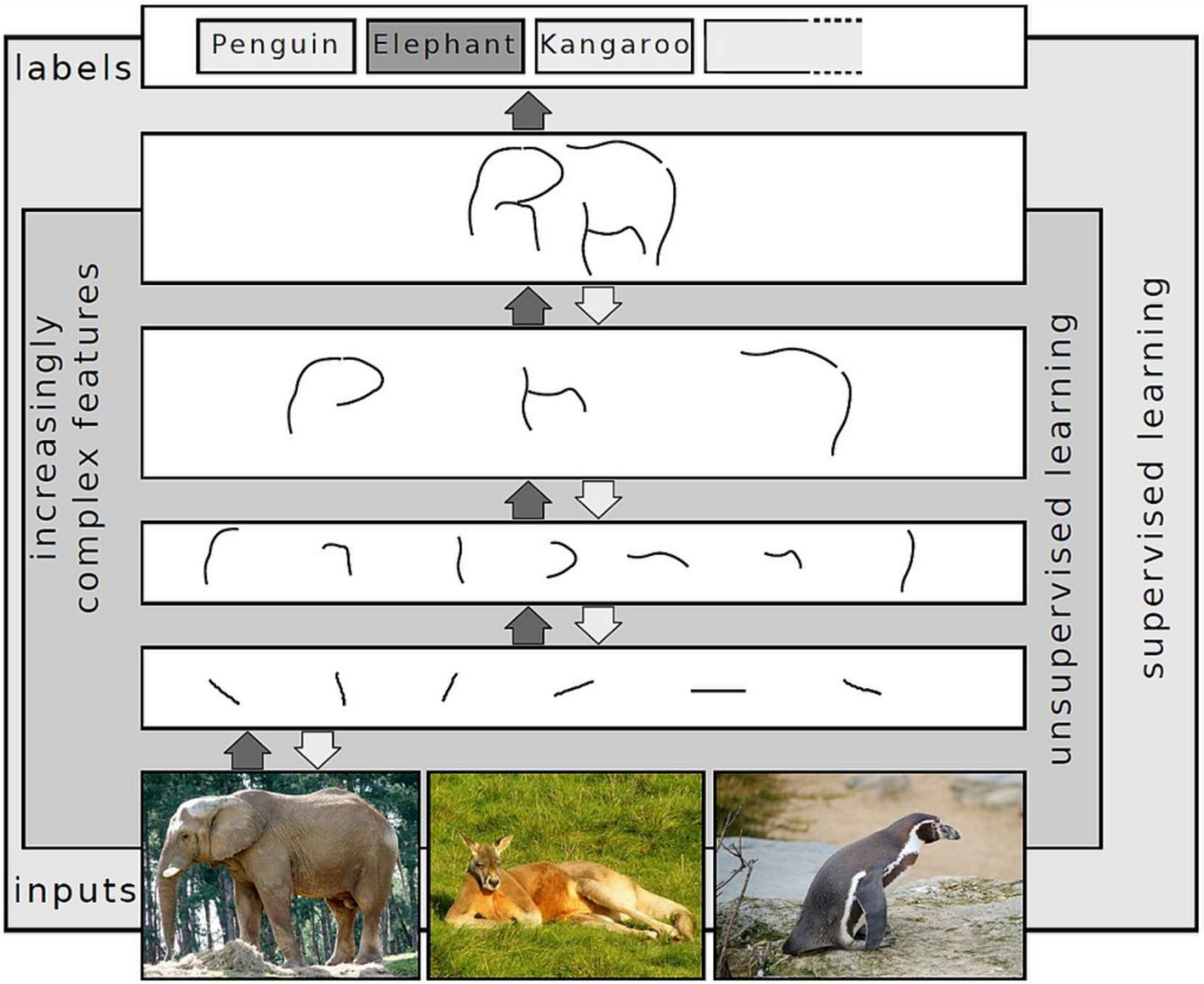
Source: https://www.quora.com/What-is-the-difference-between-deep-learning-and-usual-machine-learning

Machine Learning vs. Deep Learning



Source: https://www.intel.com/content/www/us/en/artificial-intelligence/posts/difference-between-ai-machine-learning-deep-learning.html

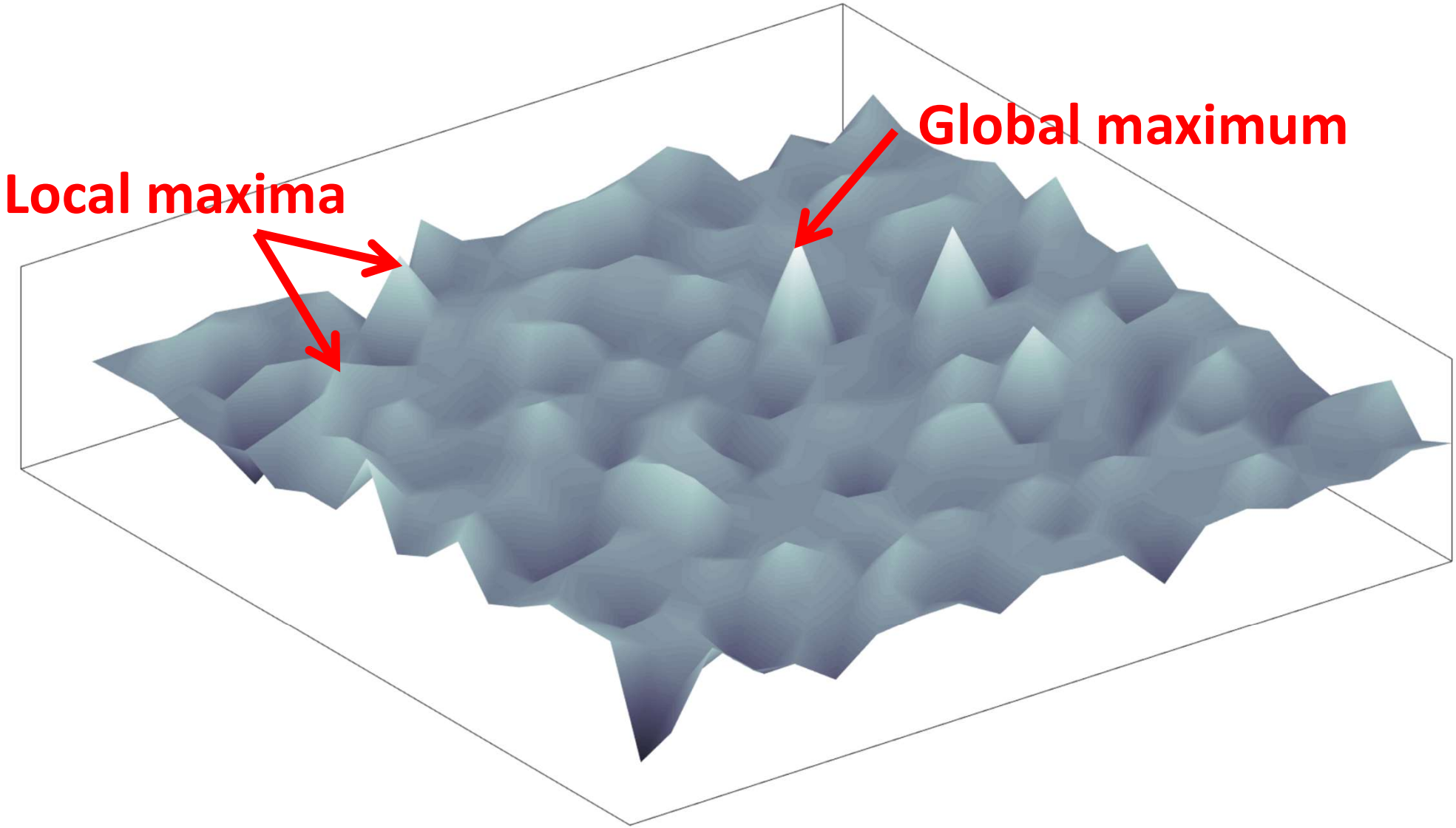
Deep Learning: Feature Extraction



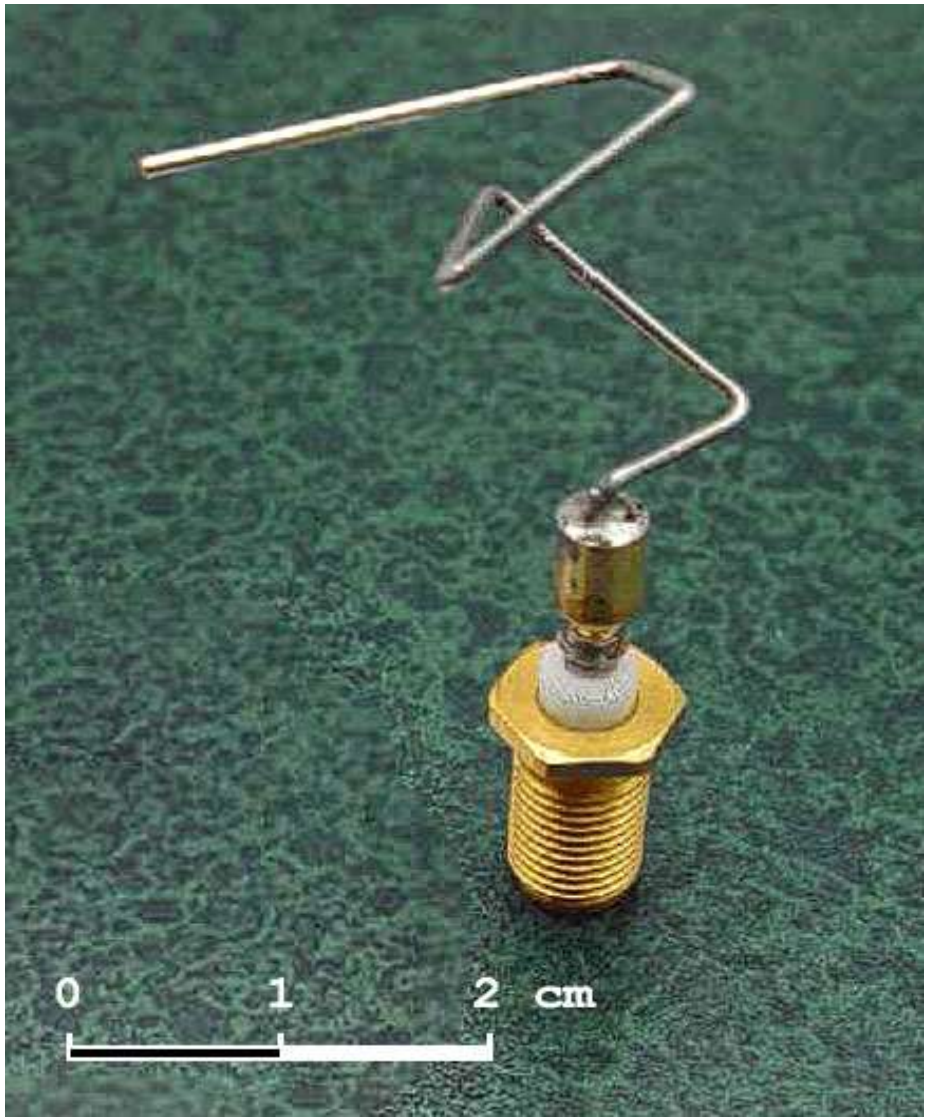
Source: https://en.wikipedia.org/wiki/Deep_learning

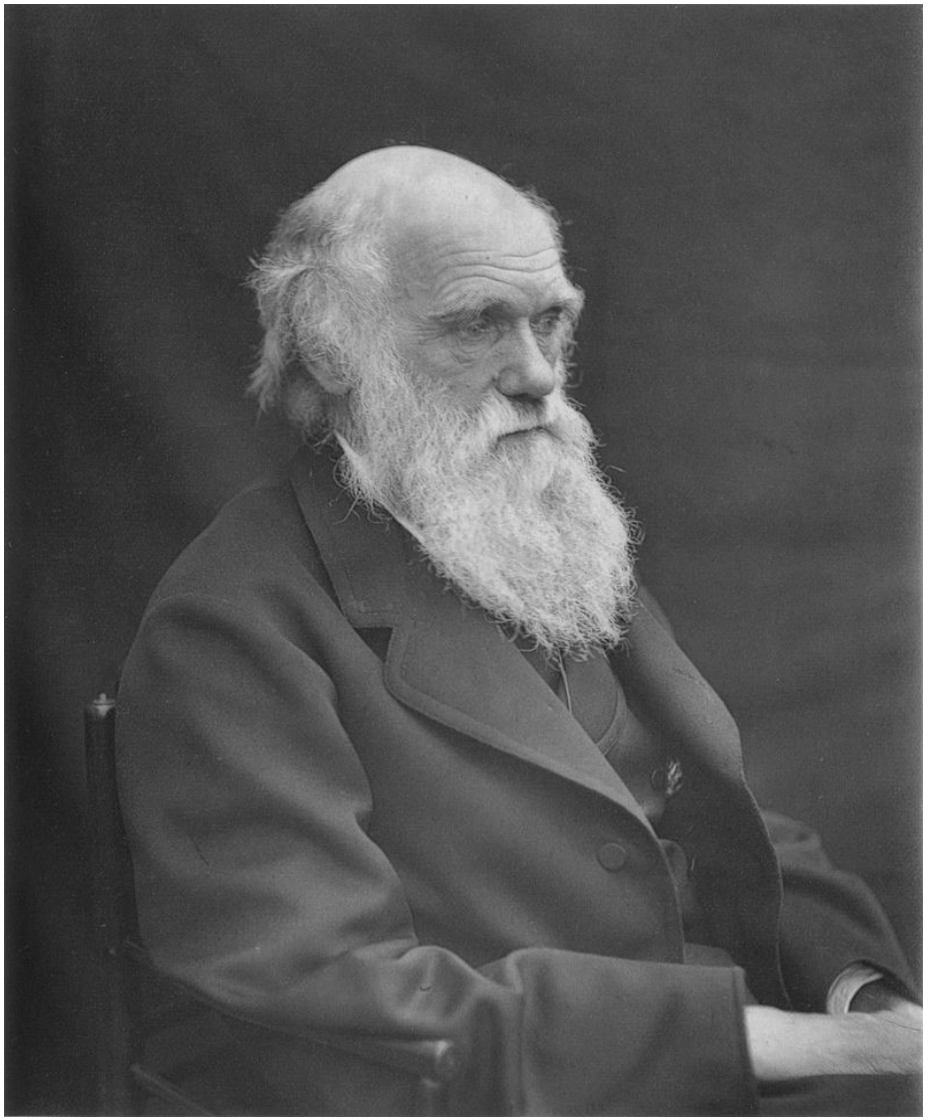
Bio-Inspired Artificial Intelligence [NOT ON FINAL EXAM]

Complex Environments



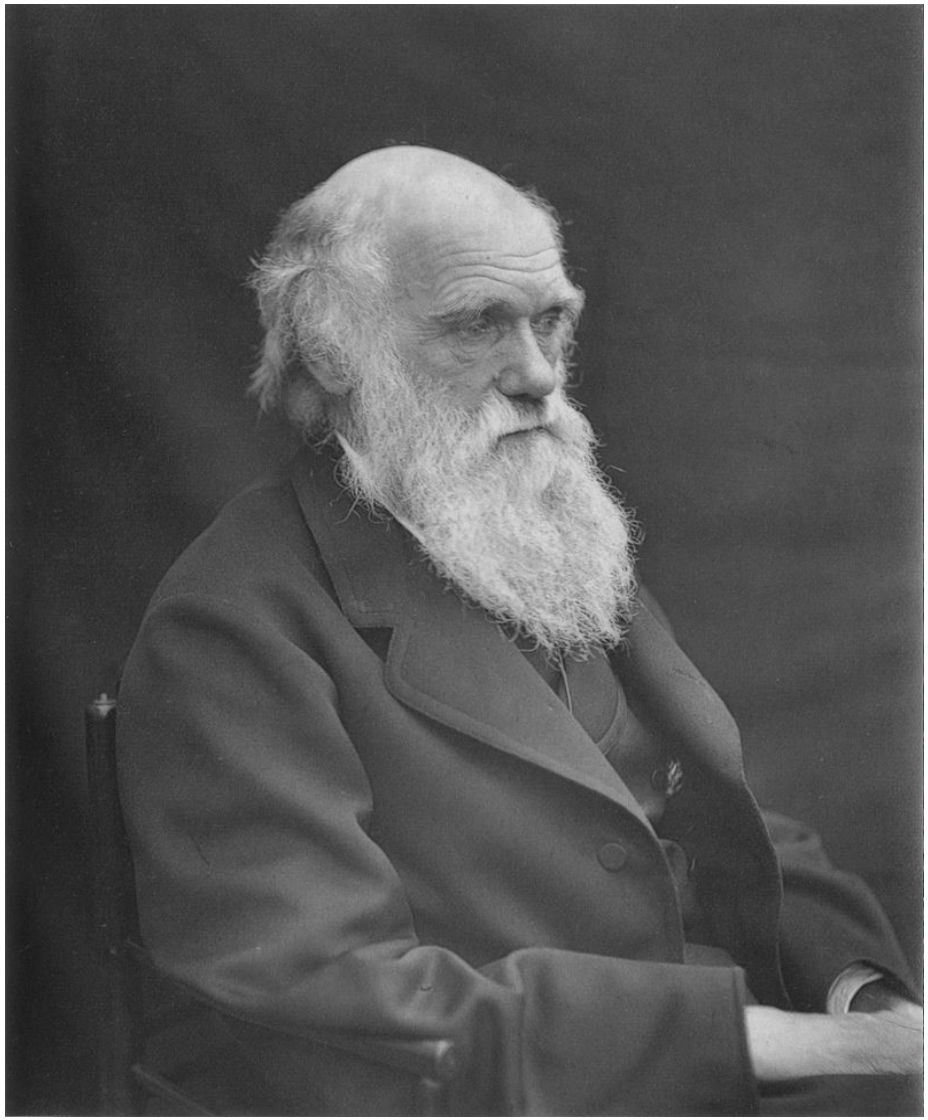
What's the Connection Here?





Source: https://wikipedia.org/

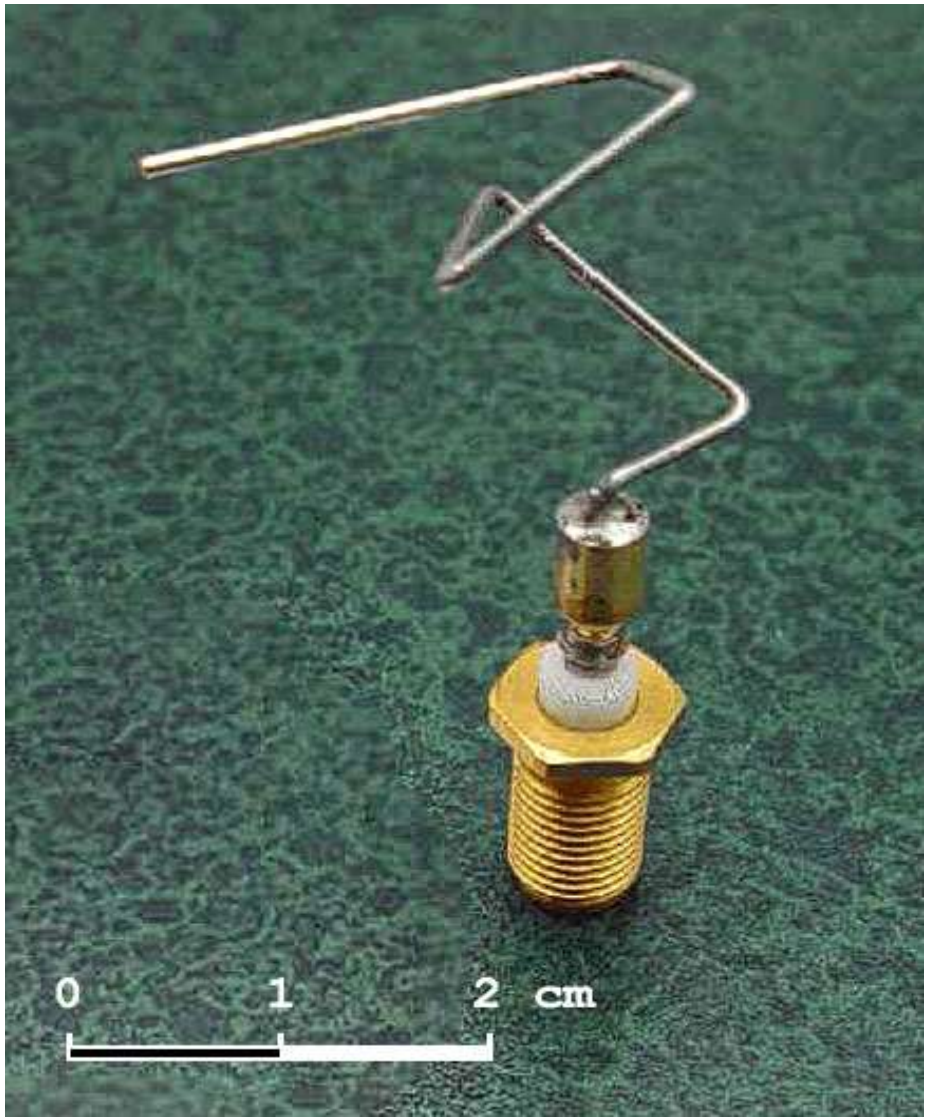
Charles Darwin



Source: https://wikipedia.org/

Charles Robert Darwin was an English naturalist, geologist and biologist, best known for his contributions to the science of evolution. His proposition that all species of life have descended over time from common ancestors is now widely accepted, and considered a foundational concept in science.

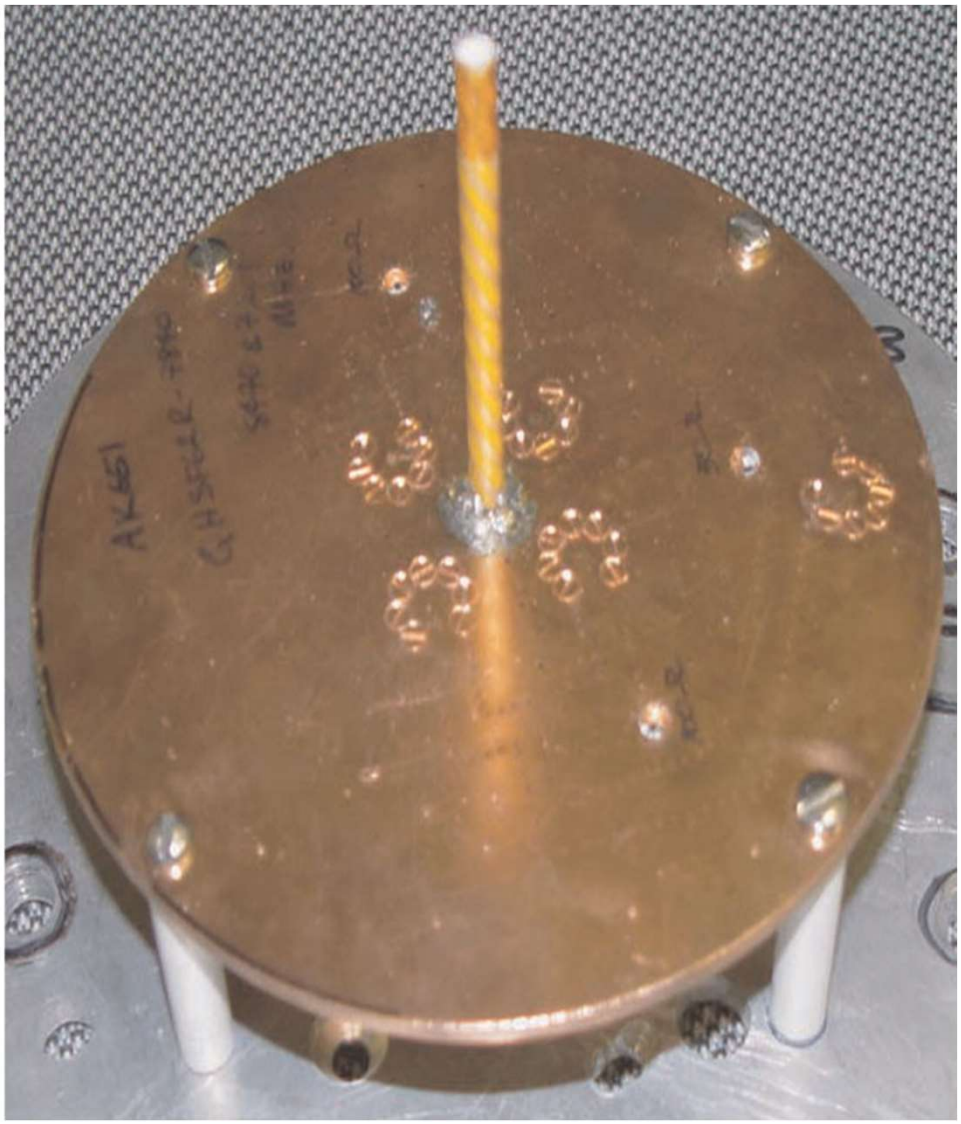
Evolved Antenna

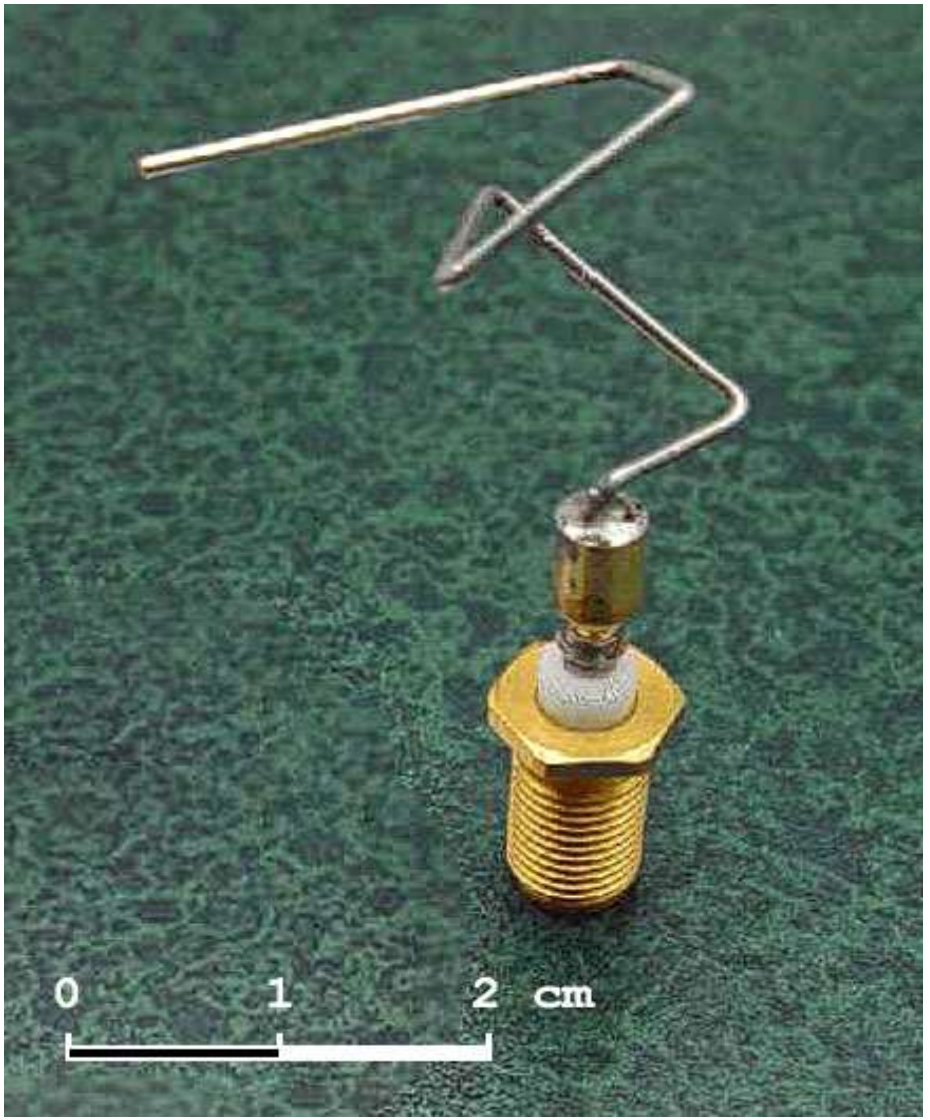


Source: https://wikipedia.org/

An evolved antenna is an antenna designed fully or substantially by an automatic computer design program that uses an evolutionary algorithm that mimics Darwinian evolution.

Engineered vs. Evolved Antenna





Source: Jason D. Lohn, Gregory S. Hornby, and Derek S. Linden - "Human-competitive evolved antennas"

Evolutionary Algorithms [Wikipedia]

An evolutionary algorithm (EA) is a subset of evolutionary computation, a generic population-based metaheuristic optimization algorithm.

An EA uses mechanisms inspired by biological evolution, such as reproduction, mutation, recombination, and selection. Candidate solutions to the optimization problem play the role of individuals in a population, and the fitness function determines the quality of the solutions (see also loss function).

Evolution of the population then takes place after the repeated application of the above operators.

Biology and Evolutionary Algorithms Background

Chromosome

A chromosome is a package of DNA with part or all of the genetic material of an organism (source: Wikipedia).

It contains genes responsible for specific traits.

Artificial Chromosome

In Evolutionary Algorithms an artificial chromosome is a genetic representation of the task to be solved.

Typically:

1 individual = 1 chromosome = 1 solution

Also called a genotype.

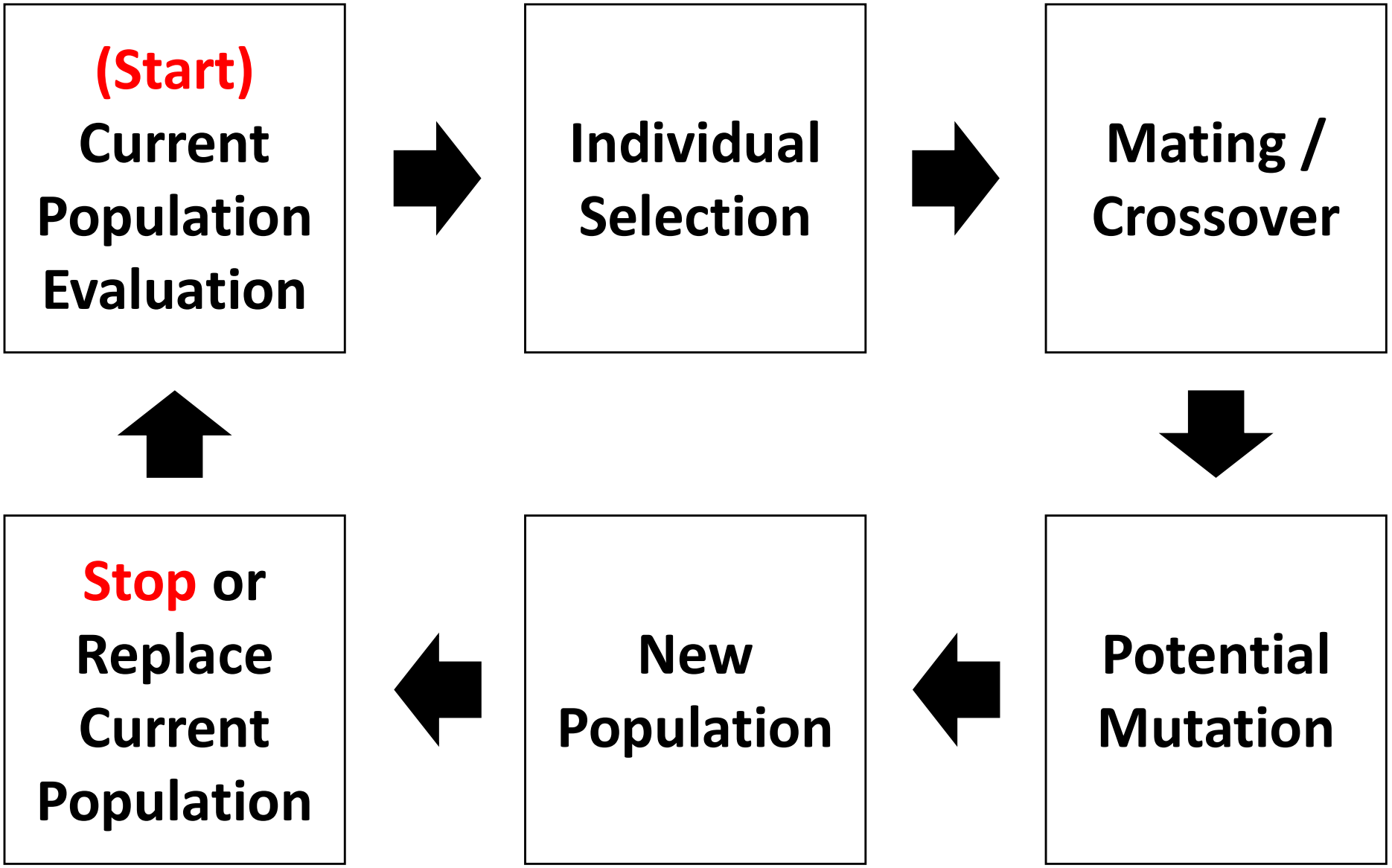
Genetic Algorithm: Roots

Directed search algorithms based on the concept of biological evolution

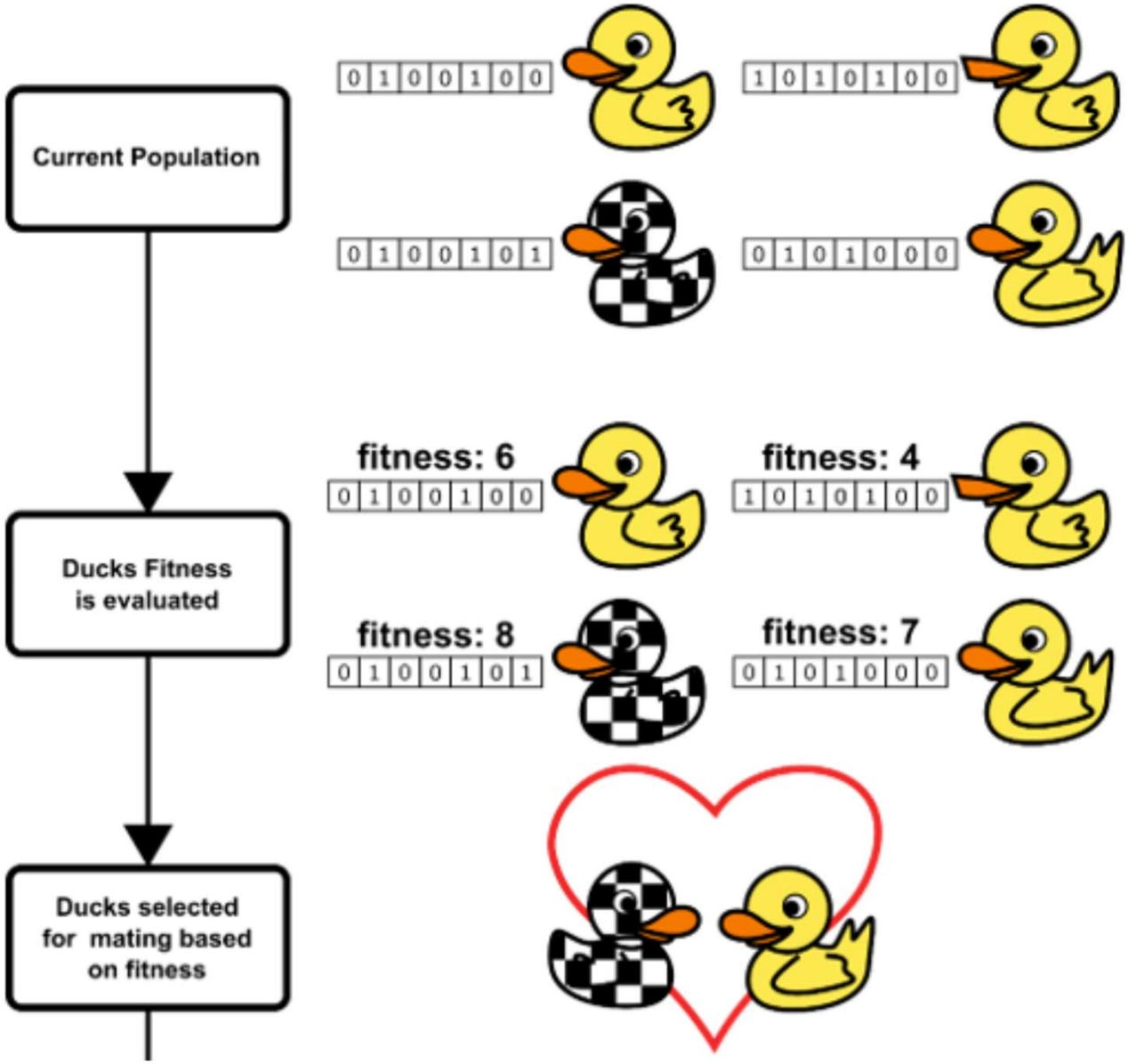
Developed by John Holland, University of Michigan (1970's)

- **to understand the adaptive processes of natural systems**
- **to design artificial systems software that retains the robustness of natural systems**

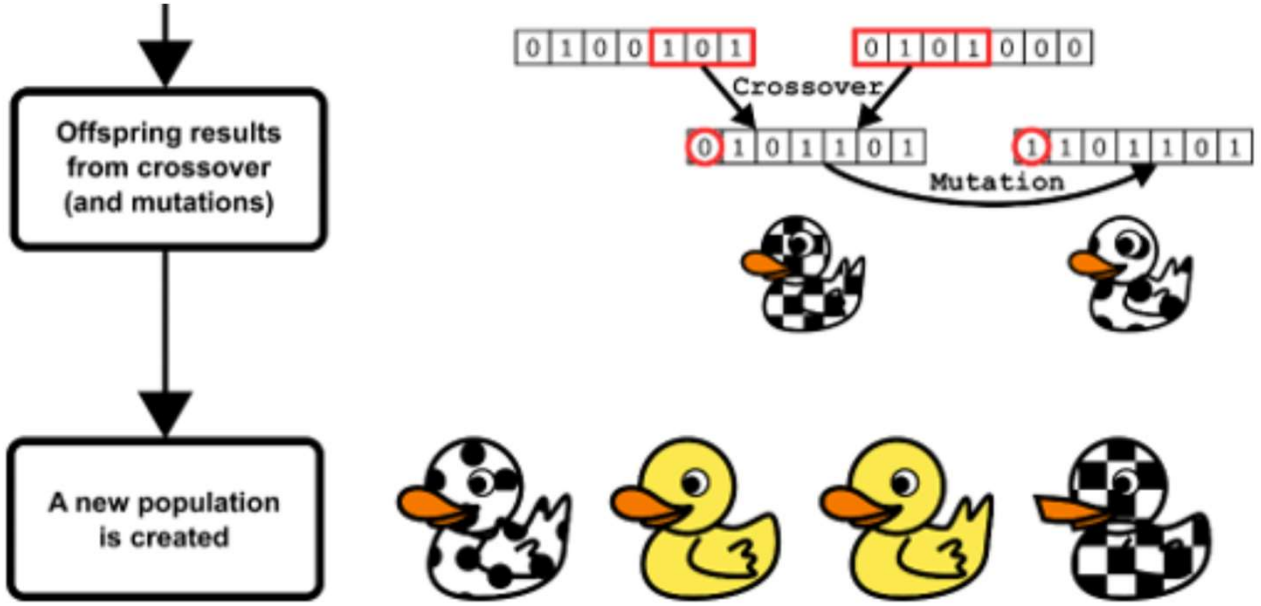
Genetic Algorithm: Process



Genetic Algorithm: The Idea



Genetic Algorithm: The Idea



Source: https://livebook.manning.com/book/algorithms-and-data-structures-in-action/chapter-18/v-14/102

Genetic Algorithm: Example

Population of points (solutions)

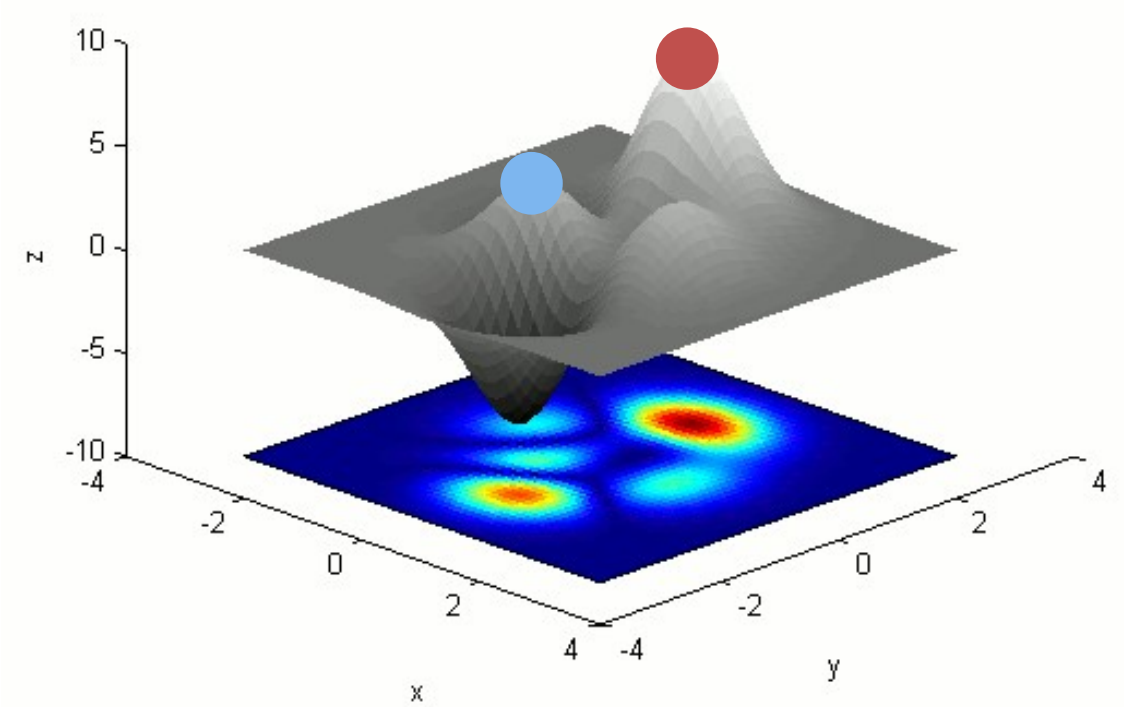
x			y		
0	1	0	1	0	1

x			y		
0	1	1	1	0	1

x			y		
0	0	0	1	0	1

x			y		
0	1	0	1	0	0

x			y		
0	0	1	0	0	0

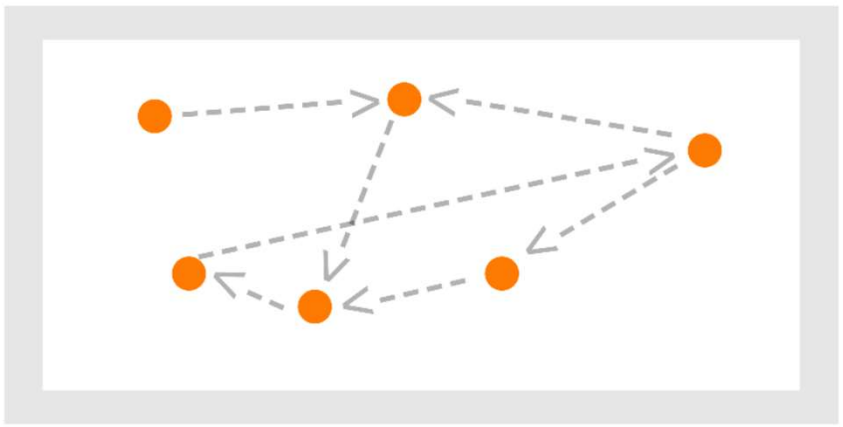


y = f(x,y) - fitness function

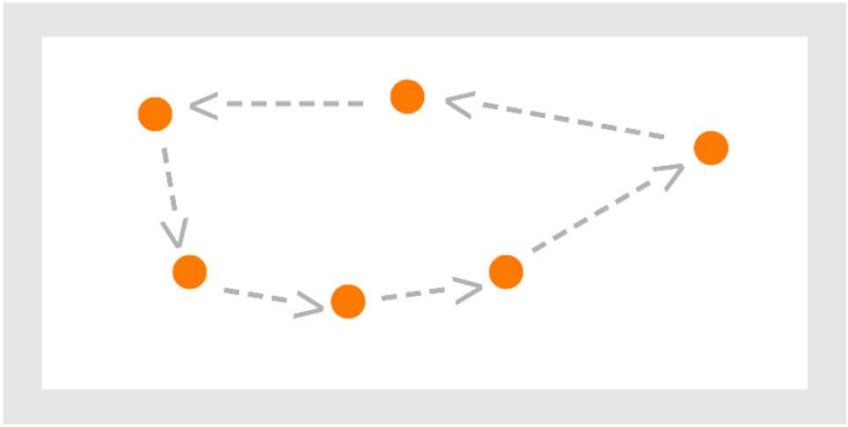
"Good enough" / local maximum

Best / global maximum

Traveling Salesman Problem



FROM THIS



TO THIS

A traveler needs to visit all the cities from a list, where distances between all the cities are known and each city should be visited just once. What is the shortest possible route that he visits each city exactly once and returns to the origin city?

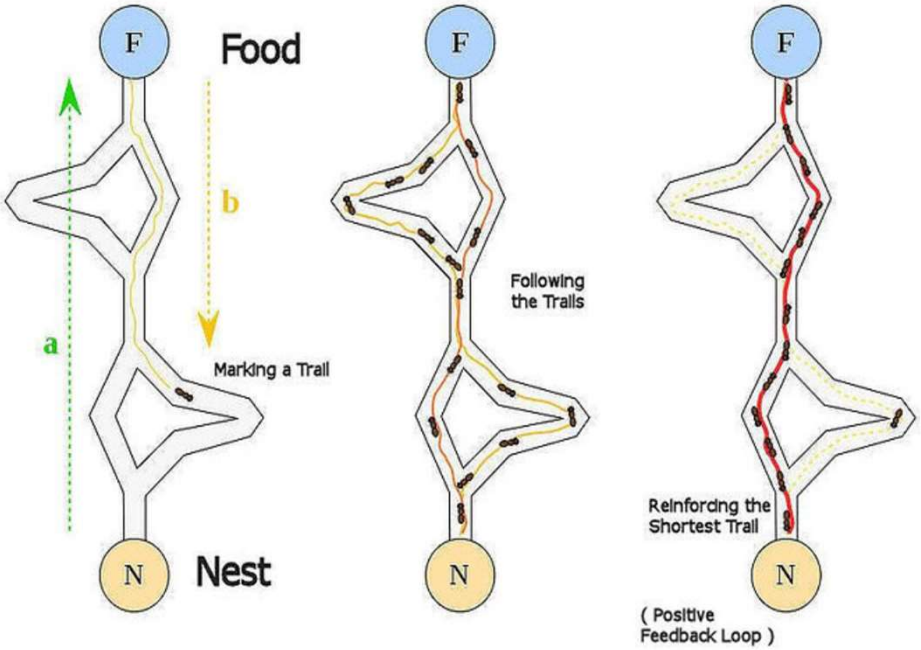
N cities → (N-1)!/2 paths | 15 cities → 43589145600 paths

Source: https://medium.com/ivymobility-developers/traveling-salesman-problem-9ab623c88fab

Example: Genetic Algorithm

http://ostap0207.github.io/web-ga-tsp/

Ant Colony Optimization: The Idea



Source: https://wikipedia.org/

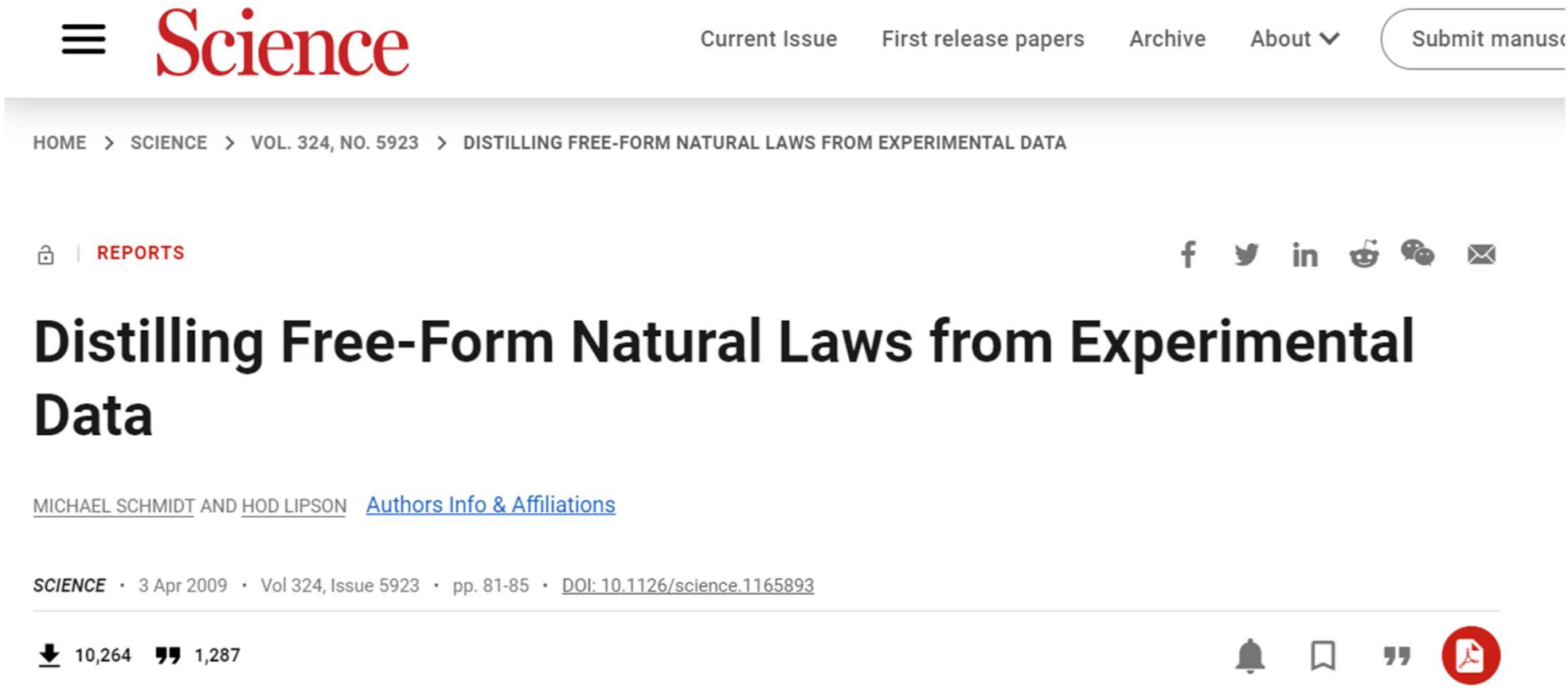
Example: Ant Colony Optimization

https://courses.cs.ut.ee/demos/visual-aco/

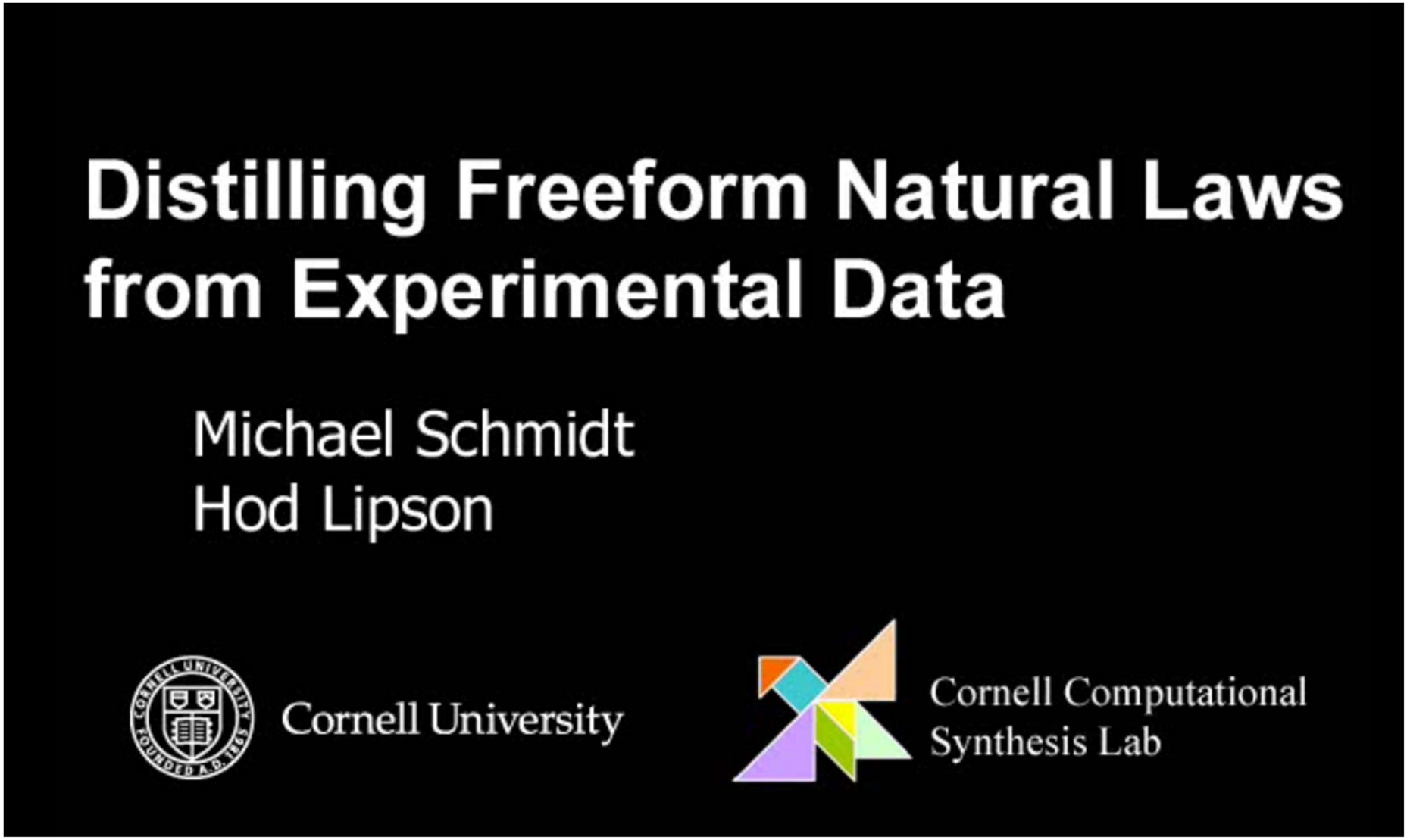
Genetic Algorithm in Action

Source: https://www.youtube.com/watch?v=qv6UVOQ0F44

Genetic Programming: Application



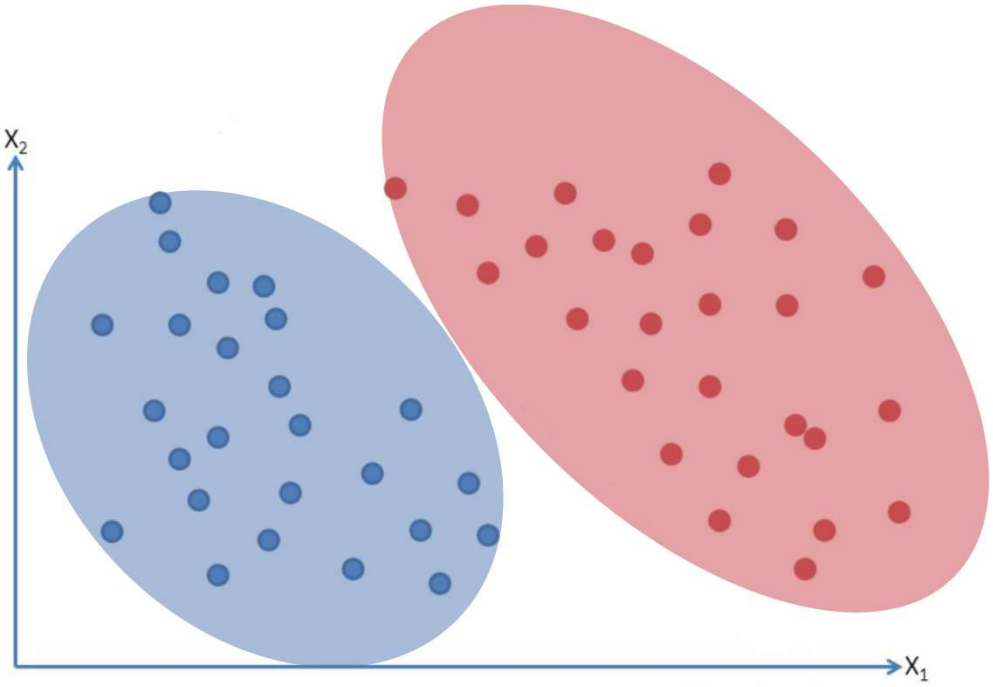
Genetic Programming: Video



Generative AI Models
[NOT ON THE FINAL EXAM]

Generative vs Discriminative Models

Generative

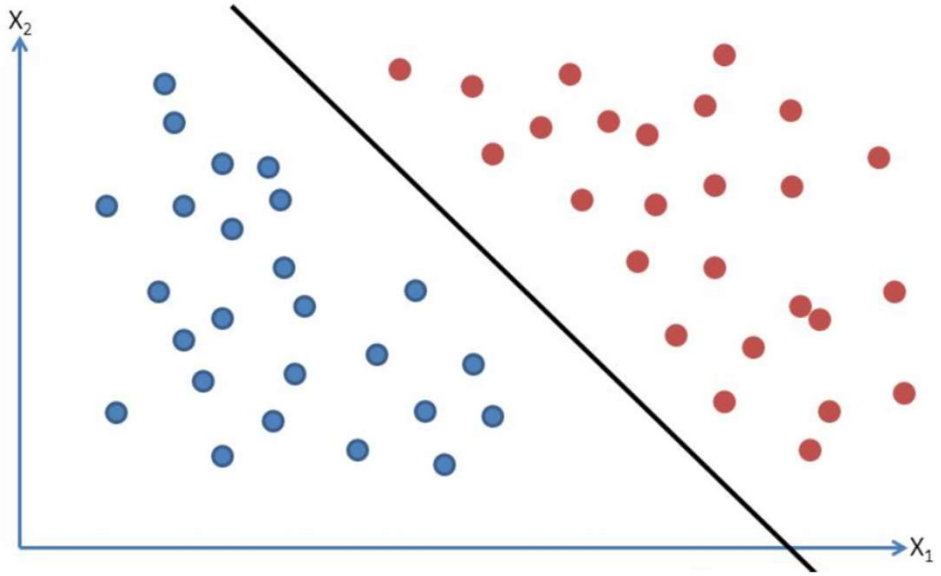


Generative model models actual distributions for EACH CLASS / LABEL / TAG

to

make a P(class | sample) **prediction**

Discriminative



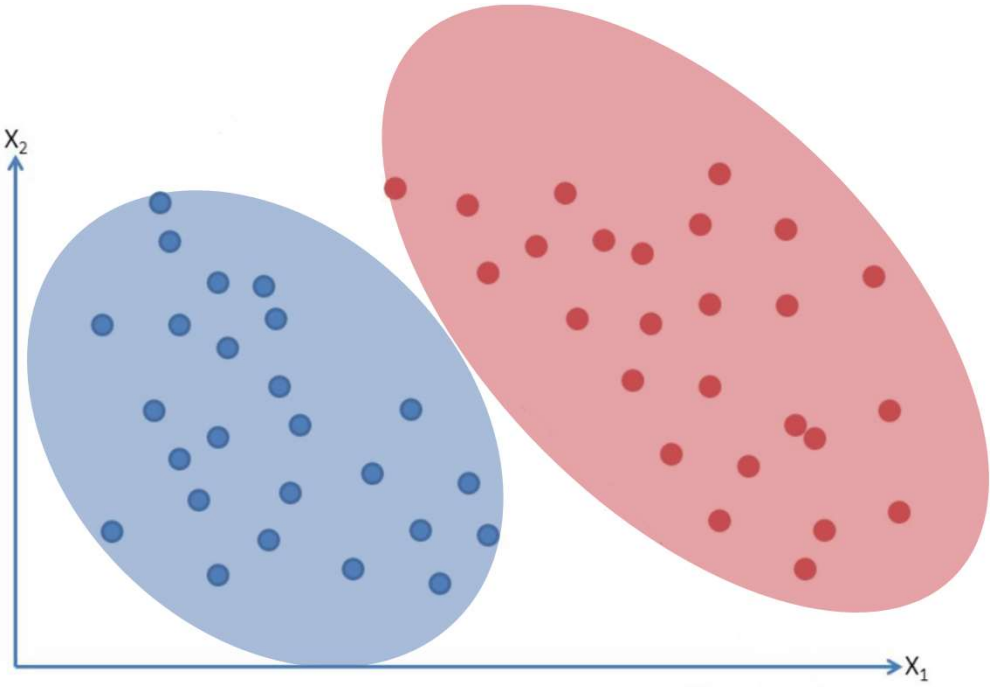
Discriminative model models the decision boundary between CLASSES / LABELS / TAGS

to

make a P(class | sample) **prediction**

Generative vs Discriminative Models

Generative

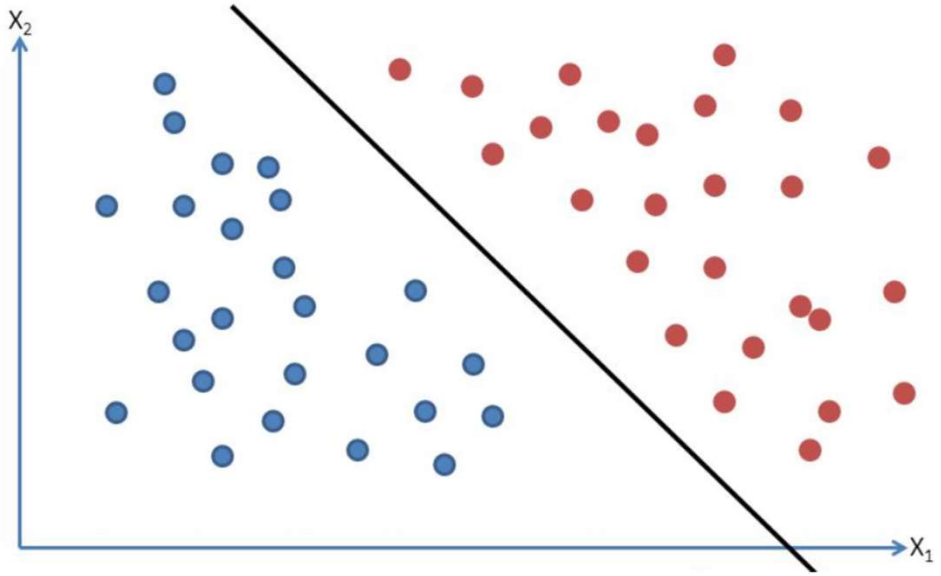


Generative model uses training data to learn P(sample, class) **joint probabilities**

and then

uses Bayes Theorem to get the P(class | sample) **prediction**

Discriminative



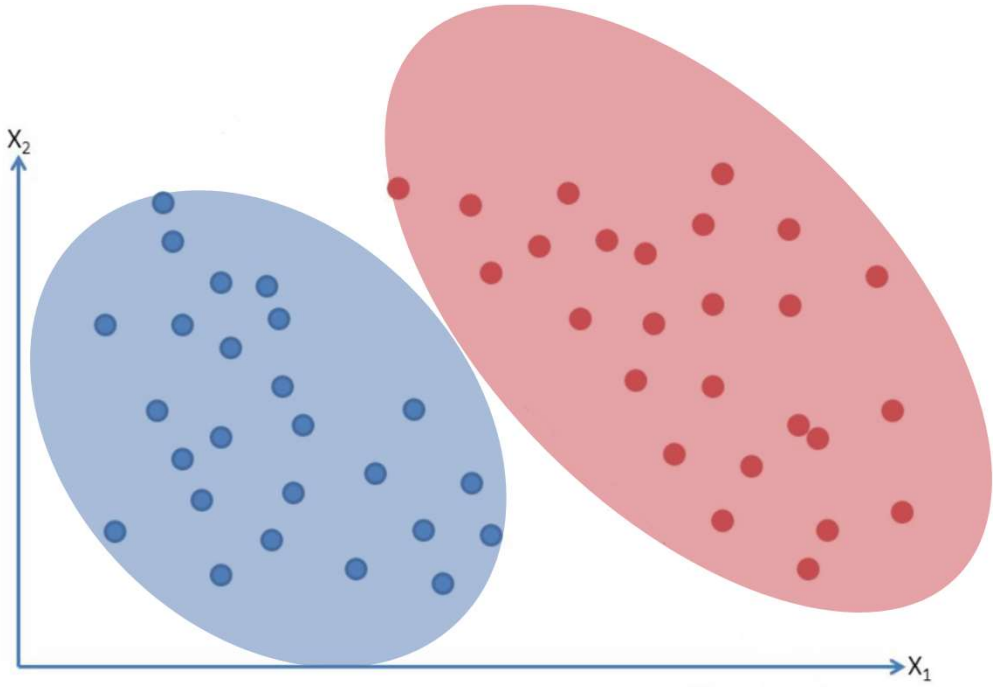
Discriminative model uses training data to learn P(class | sample) **conditional probability**

and then

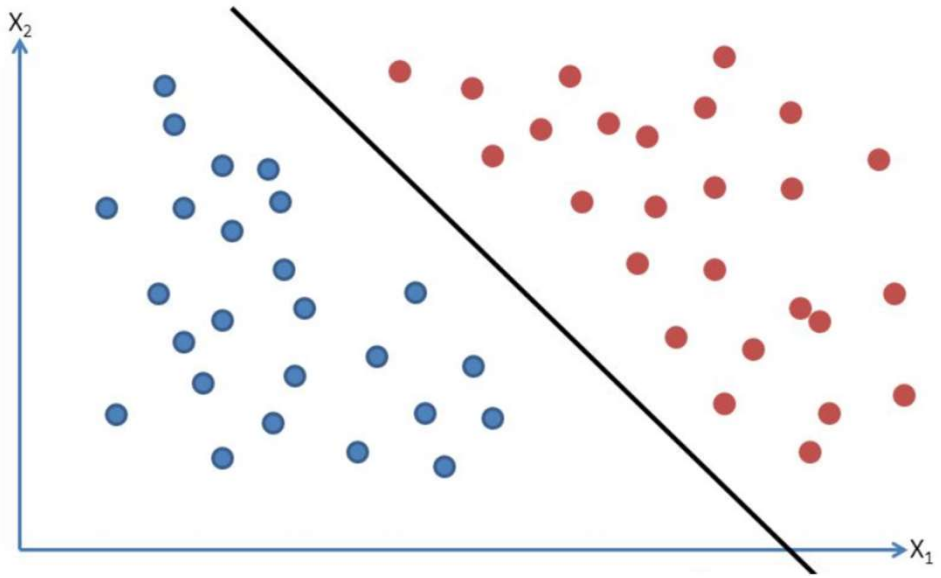
uses it to make a prediction

Generative vs Discriminative Classifier

Generative



Discriminative



Generative classifiers:

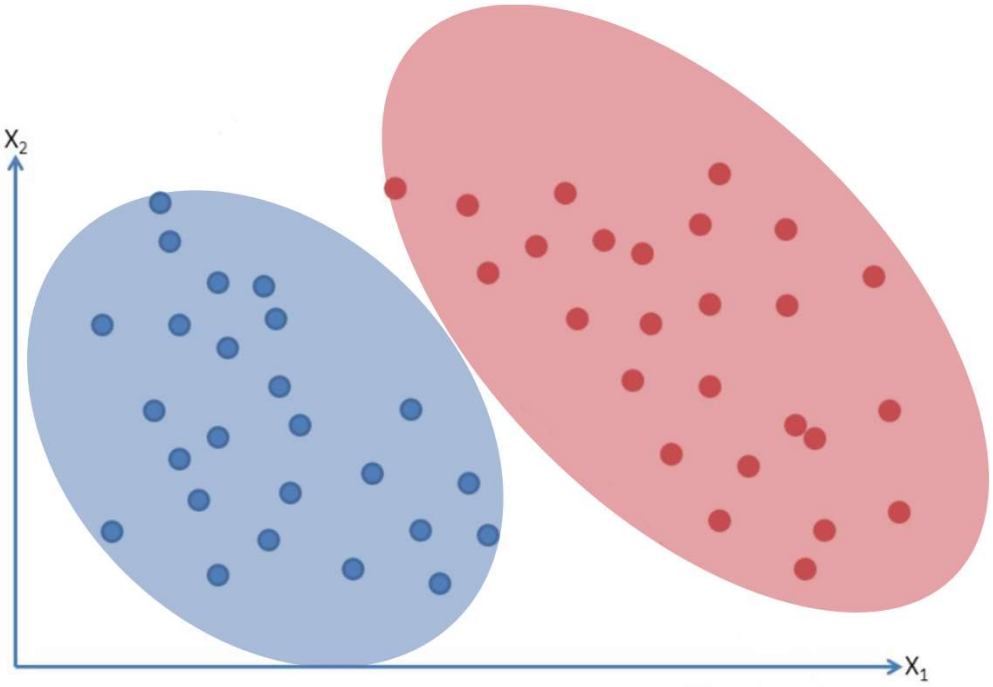
- **Assume some form of** P(class), P(sample | class)
- **Estimate** P(class), P(sample | class) **using training data**
- **Use Bayes Theorem to calculate** P(class | sample)

Discriminative classifiers:

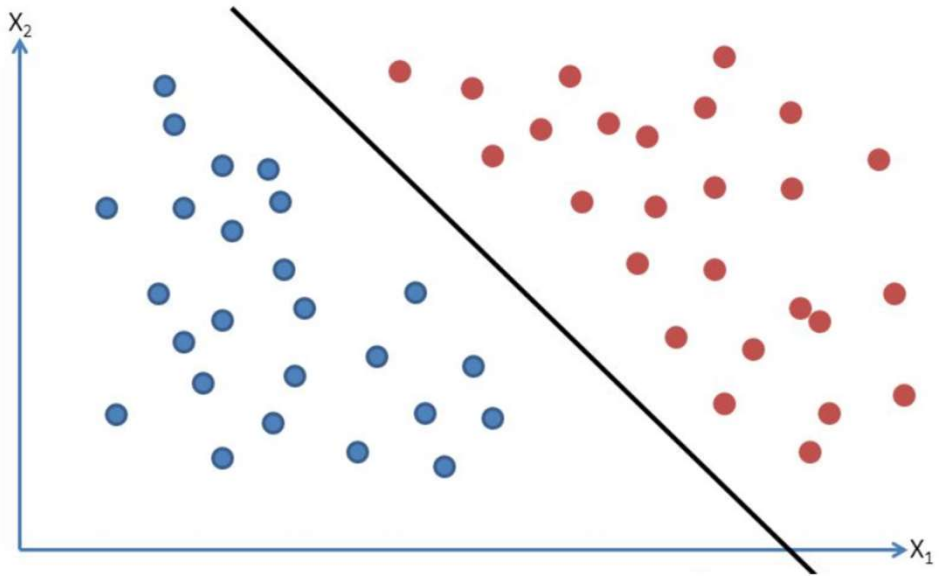
- **Assume some form of** P(class | sample)
- **Estimate** P(class | sample) **using training data**

Generative vs Discriminative Classifier

Generative



Discriminative



Generative classifiers:

- Naive Bayes
- Bayesian networks
- Markov random fields
- Hidden Markov Models (HMM)

Discriminative classifiers:

- Logistic regression
- Support Vector Machines
- Traditional neural networks
- k-Nearest Neighbors
- Conditional Random Fields (CRF)s

Generative AI Model: the Idea

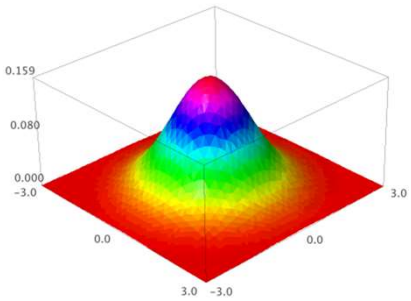
Reality: some sample space (possibly infinite)

Sample

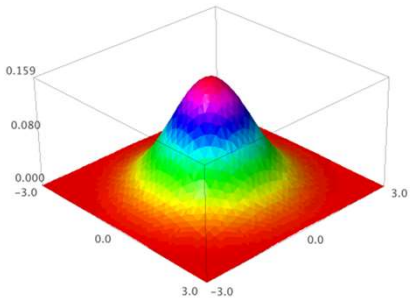
Dataset: Subset of "Reality" (finite)

New sample

Model



Estimate



Generate / sample

Probability distribution P_{real}

Probability distribution P_{model}

Generative AI Model: the Idea

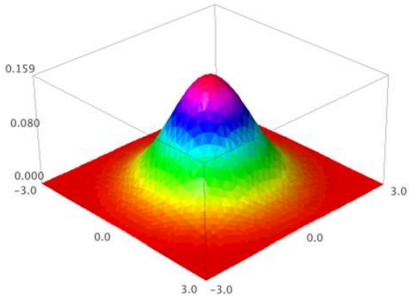
Reality:
All Possible cat faces

Sample

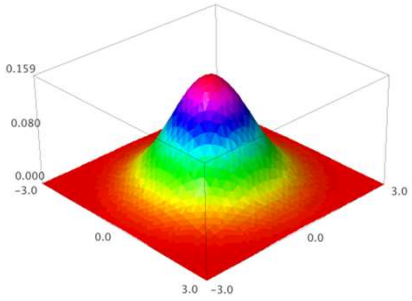
Dataset:

New sample:

Model



Estimate



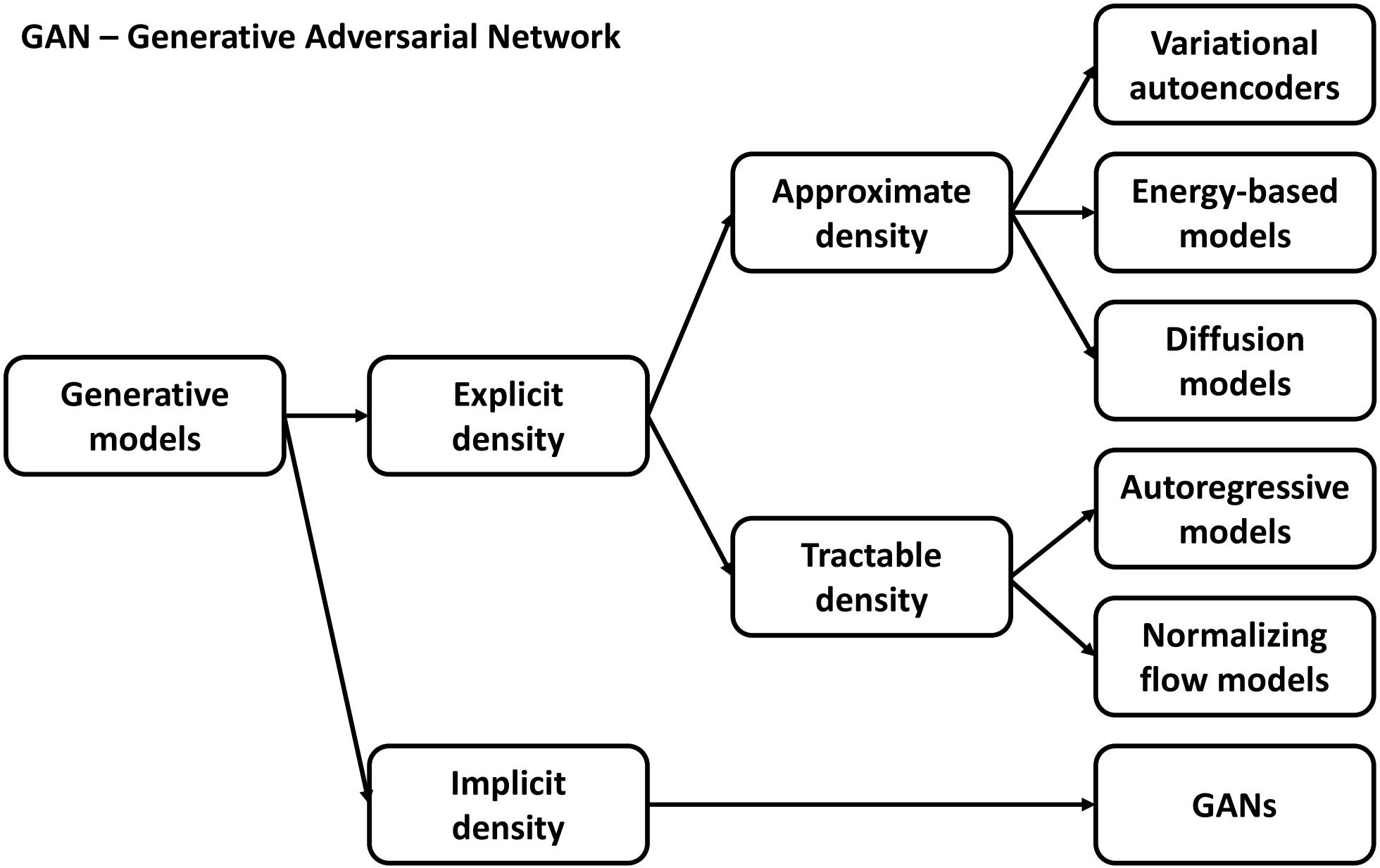
Generate / sample

Probability distribution P_{real}

Probability distribution P_{model}

Taxonomy of Generative AI Models

GAN – Generative Adversarial Network



Explicit Density

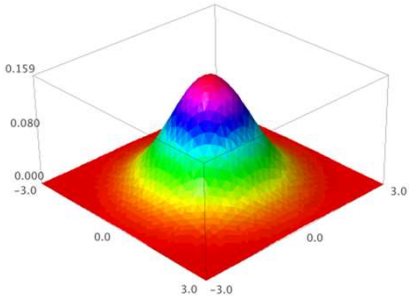
Reality: All Possible cat faces

Sample

Dataset:

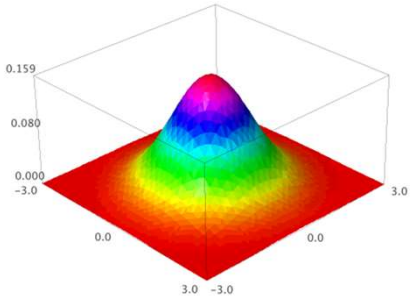
New sample:

Model



Probability distribution P_{real}

Estimate



Probability distribution P_{model}

Sample

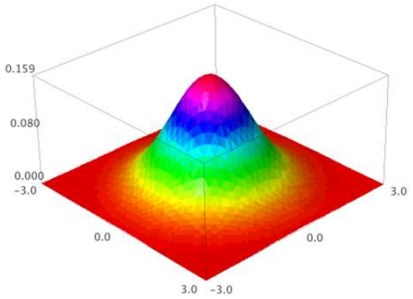
Implicit Density

Reality:
All Possible cat faces

Sample

Dataset:

New sample:



Probability distribution P_{real}

Model

Stochastic generation process

Generate

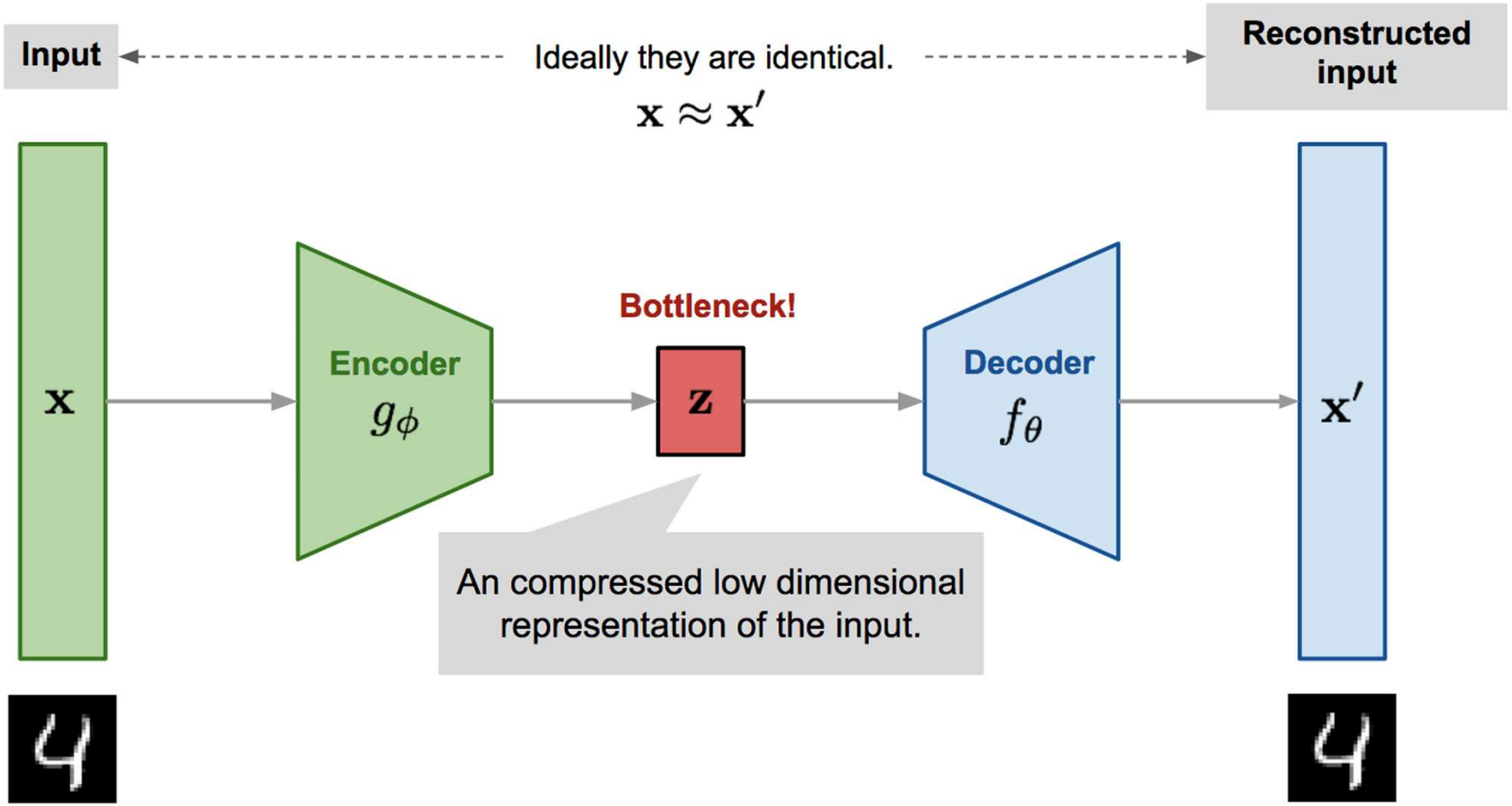
Tractable vs. Approximate Density

Tractable density models place constraints on the model architecture so that the density function has a form that makes it easy to calculate.

Approximate density models use variety of techniques to approximate the density function:

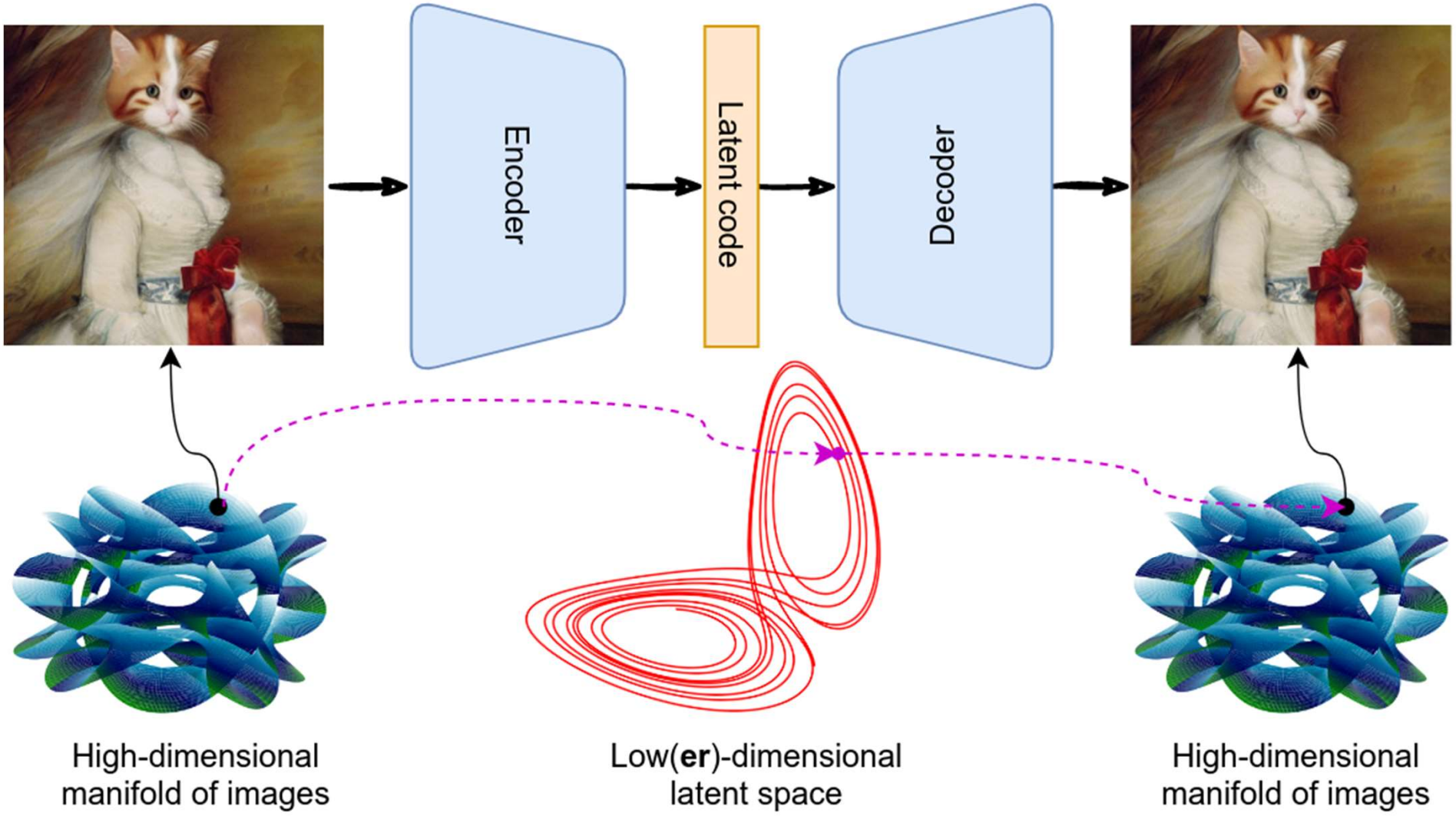
- **latent vectors**
- **denoising**

Autoencoder Model



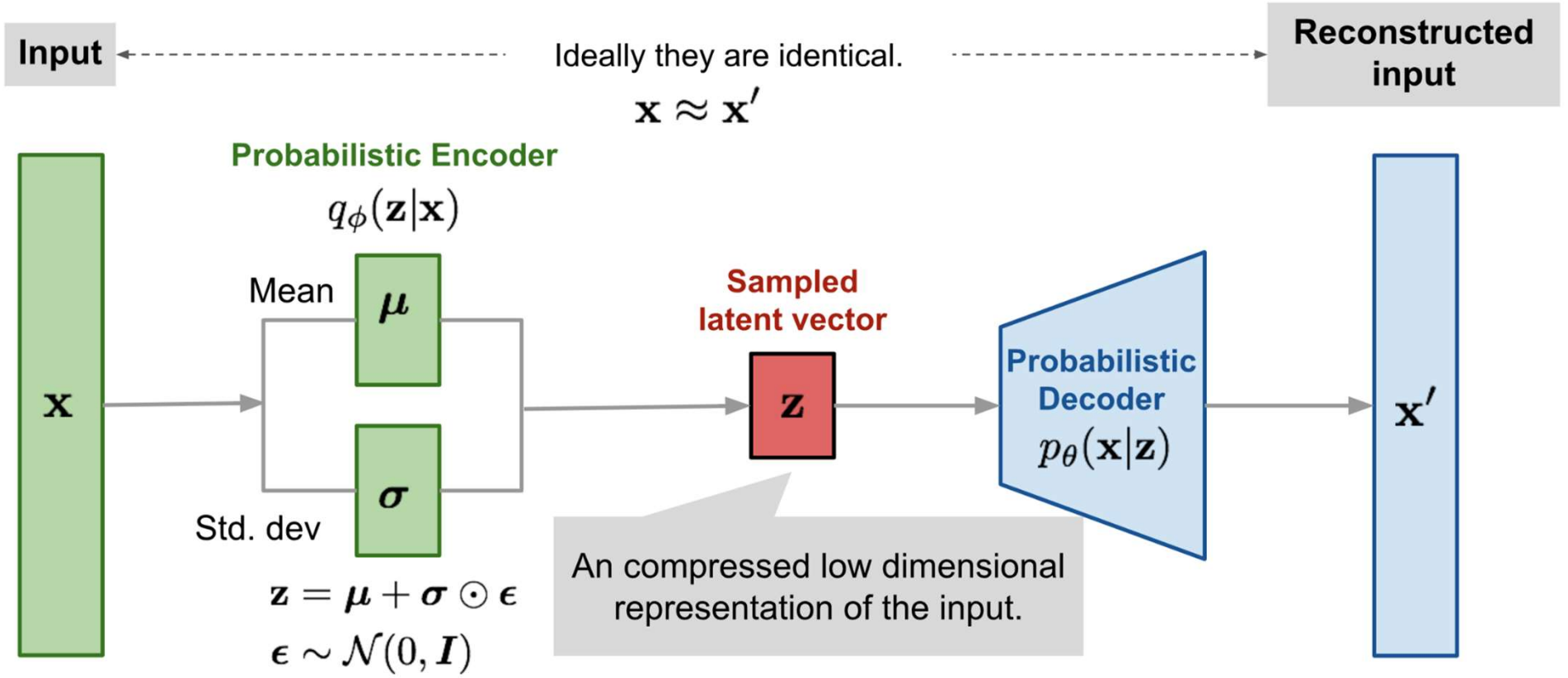
Source: https://lilianweng.github.io/posts/2018-08-12-vae/

Latent Space



Source: https://synthesis.ai/2023/03/21/generative-ai-ii-discrete-latent-spaces/

Variational Autoencoder Model



Source: https://lilianweng.github.io/posts/2018-08-12-vae/

Autoregressive Model (GPT-3)

What is it?

Generative Pre-trained Transformer 3 (GPT-3) is an **autoregressive language model that uses deep learning to produce human-like text**. It is the third-generation language prediction model in the GPT-n series (and the successor to GPT-2) created by OpenAI, a San Francisco-based artificial intelligence research laboratory.

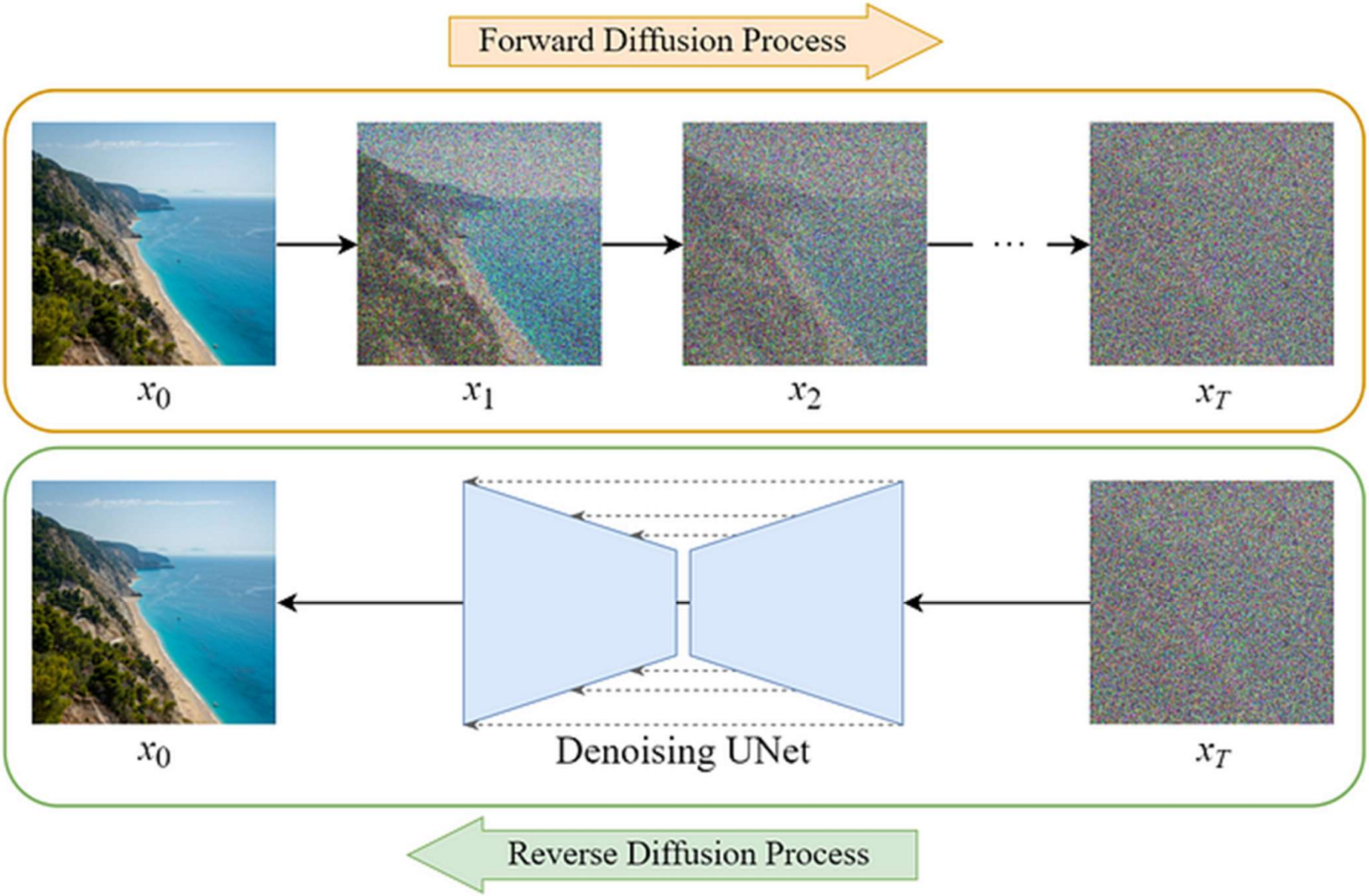
Size:

175 billion machine learning parameters

~45 GB

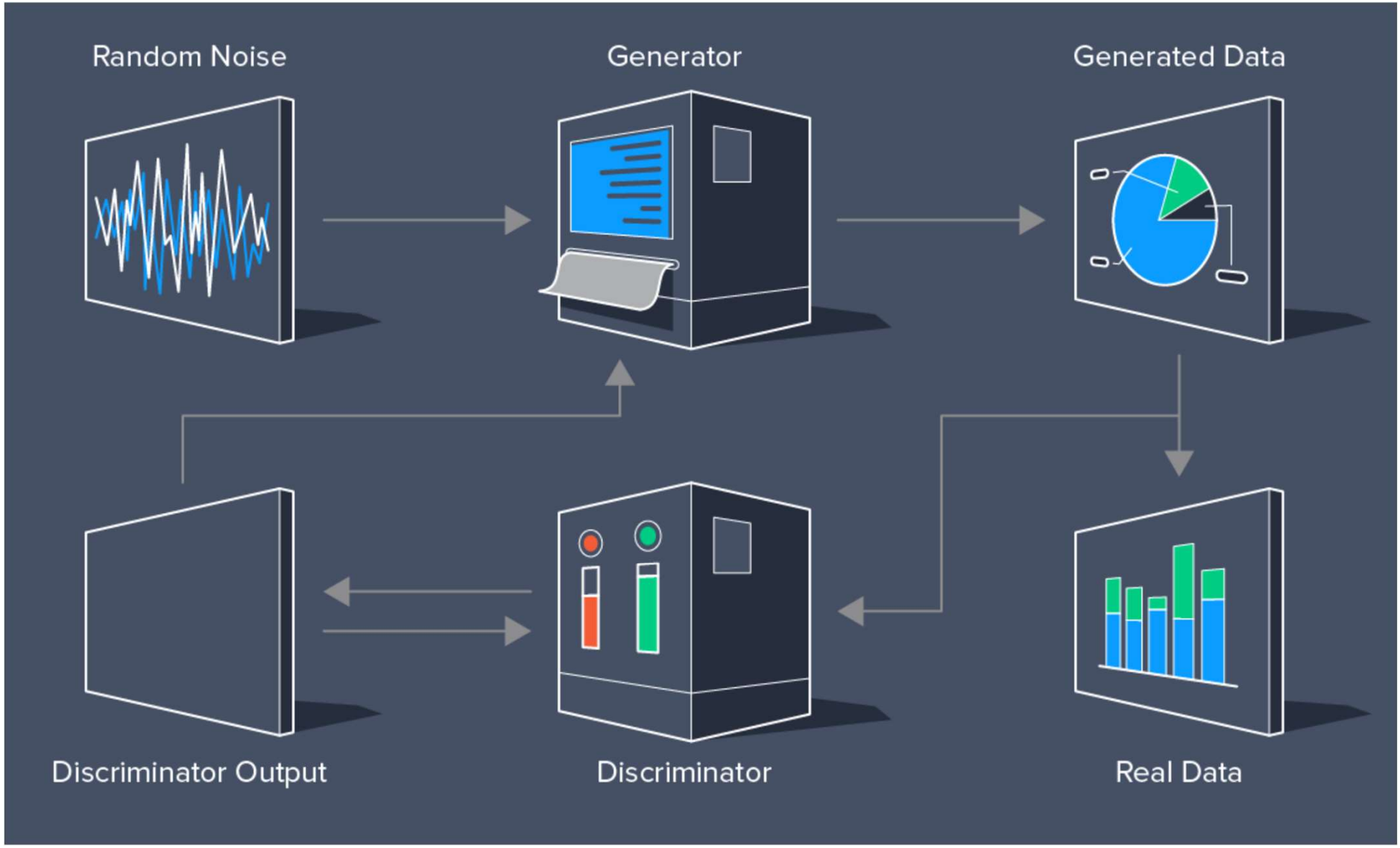
Source: Wikipedia

Diffusion Model



Source: https://medium.com/@steinsfu/stable-diffusion-clearly-explained-ed008044e07e

Generative Adversarial Network



Source: https://www.toptal.com/machine-learning/generative-adversarial-networks