Introduction to the Course

Generative Models: Fundamentals and Applications

Shuigeng Zhou, Wenhui Liu

- sgzhou@fudan.edu.cn
- wh_liu@fudan.edu.cn

College of Computer Science and Artificial Intelligence

2026/9/11

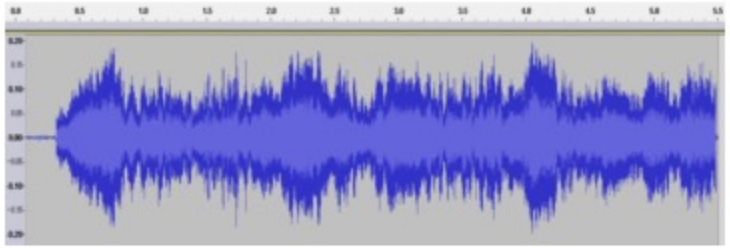
Outline

- What are generative models?
- Why generative models?
- Discriminative models vs. Generative models
- Applications of generative models
- Arrangements for the course

Generative Models Now Everywhere…

- **Challenge**
  - Understand *complex, unstructured* inputs
  - Generate *unseen* samples

Computer Vision
image/video generation, restoration, style transfer, scene understanding



Speech and Audio
text-to-speech, voice conversion, music generation, speech enhancement

Natural Language Processing
text generation, dialogue systems, machine translation, code generation

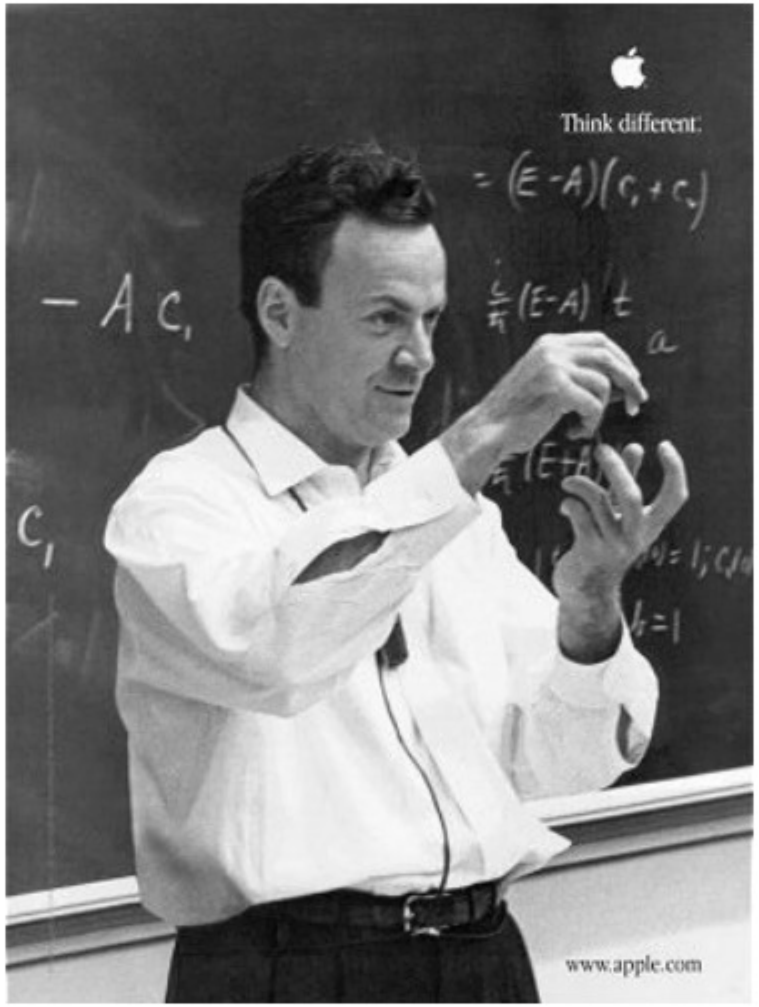
Robotics
dynamics modeling, trajectory prediction, motion planning, environment simulation

What are Generative Models?

- *“What I cannot create, I do not understand”*

    -- Richard Feynman (Physicist, 1965 Nobel Prize winner)

- Generative modeling
    - What I understand, I can create



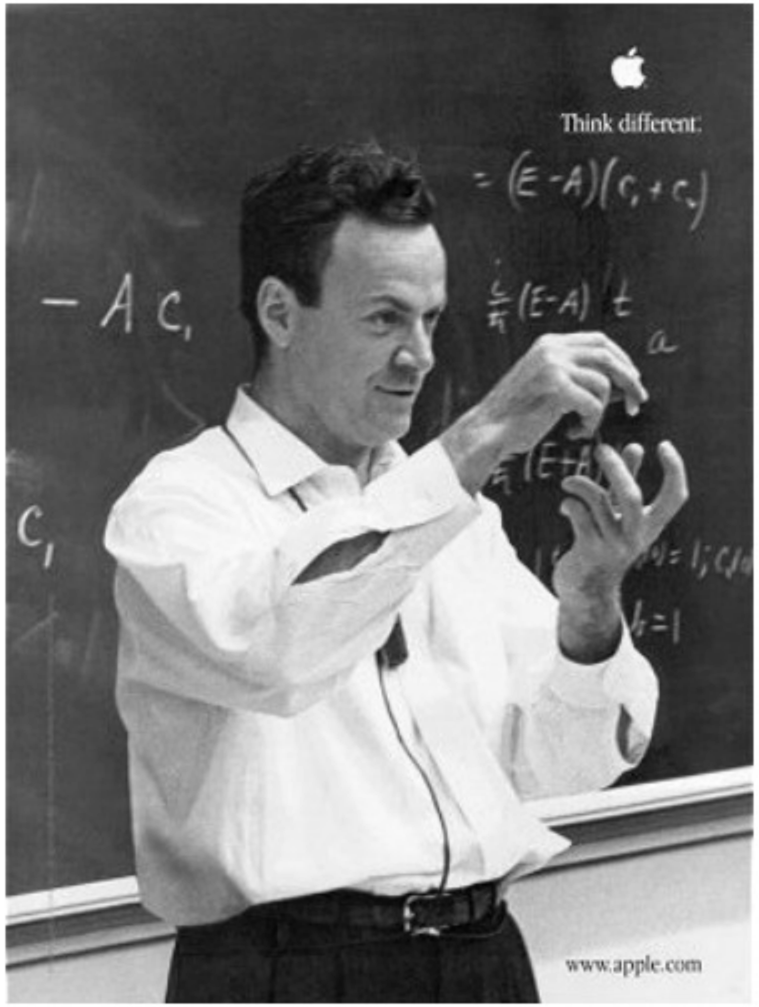
What are Generative Models?

- *"What I cannot create, I do not understand"*

  -- Richard Feynman (Physicist, 1965 Nobel Prize winner)

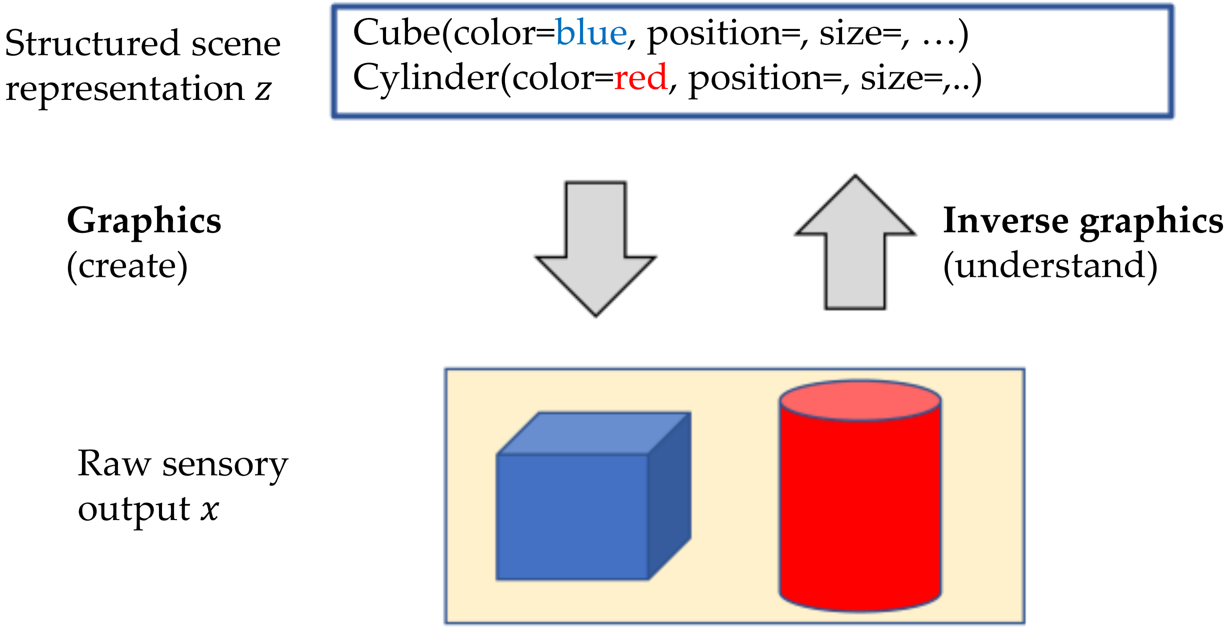
- Generative modeling
  - What I understand, I can create

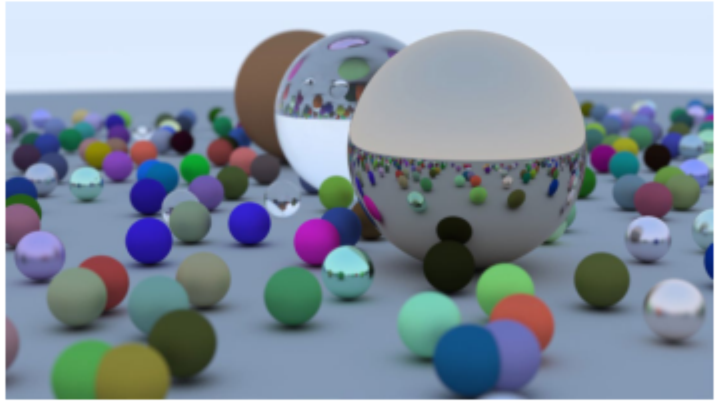
    How?
    Through a probabilistic / statistical view



Generative Modeling

- An example case: How to generate natural images with a computer?



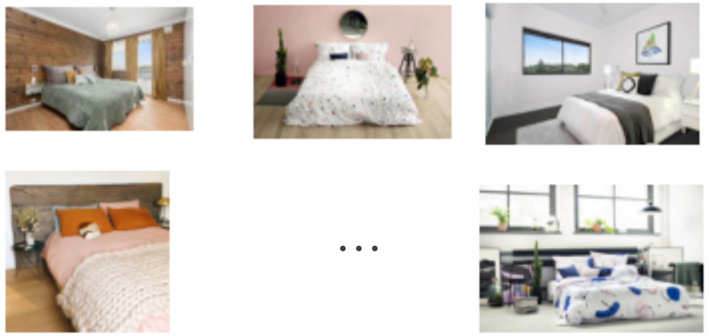


Generated balls via ray tracing
Shirley et al., 2016

- Many generative models will have similar structure (generation + inference)

Statistical Generative Models

- Statistical generative models are **learned from data**



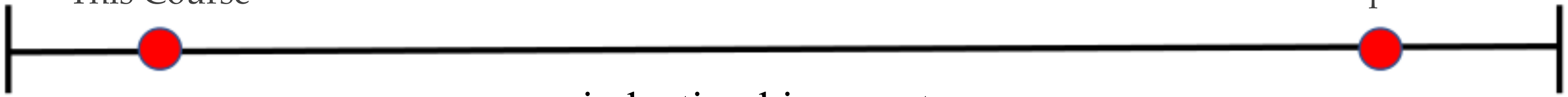
Data
(e.g., images of bedrooms)

+

Prior Knowledge
(e.g., physics, materials, ..)

- Inductive bias is unavoidable; it may appear as a prior, architecture, objective, data augmentation, or optimization choice.

This Course — Graphics

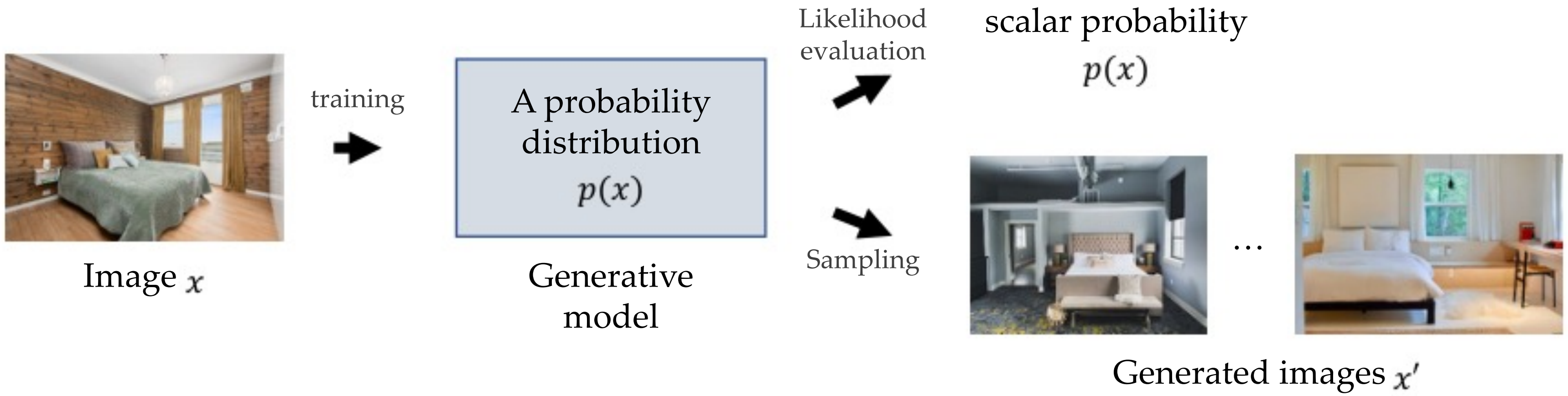
Data  Prior Knowledge

inductive-bias spectrum

* *Wolpert, D. H., & Macready, W. G, 1997. No Free Lunch Theorems for Optimization.*

Statistical Generative Models

- A statistical generative model is **a probability distribution** $p(x)$
    - **Data:** samples (e.g., images of bedrooms)
    - **Prior knowledge:** parametric form (e.g., Gaussian?), loss function (e.g., maximum likelihood?), optimization algorithm, etc.

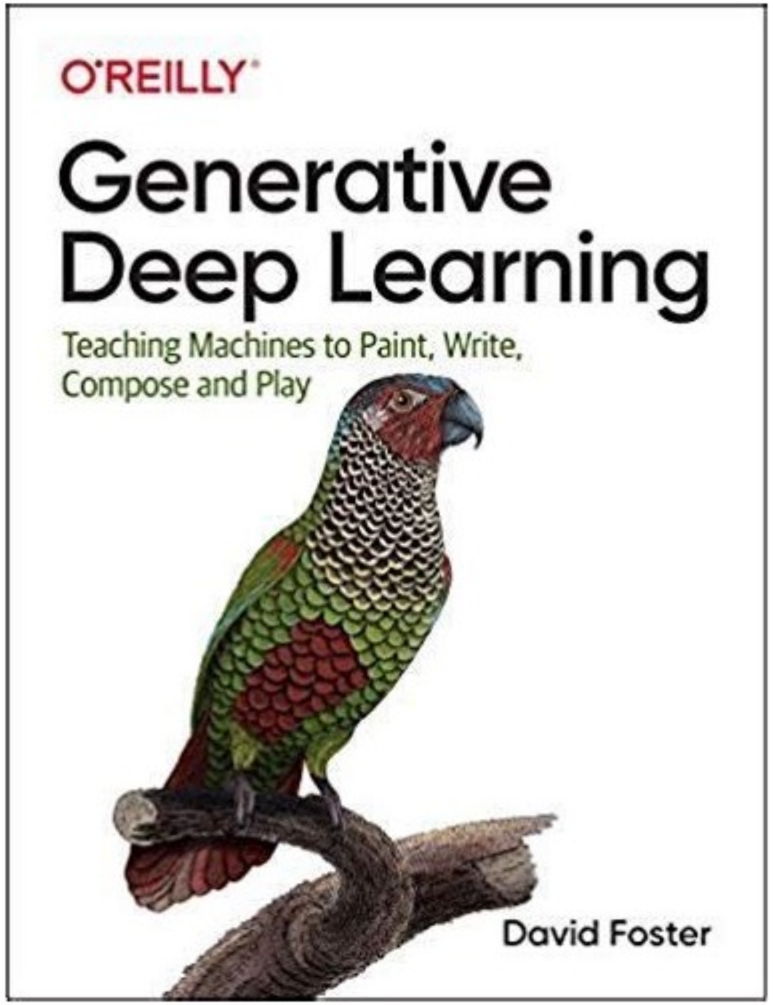


- It is generative because sampling from $p(x)$ generates new images

Statistical Generative Models

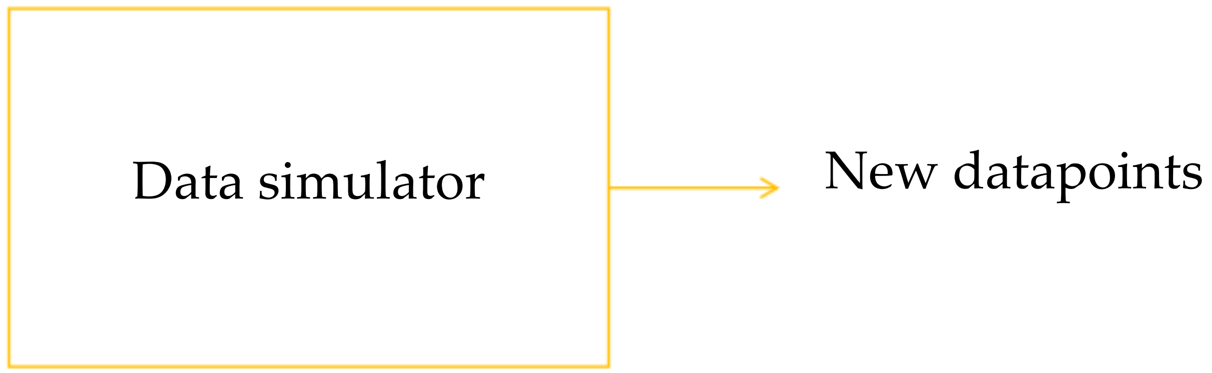
- **A broad definition**
  - *A generative model describes how a dataset is generated, in terms of a probabilistic model.*
    By sampling from this model, we are able to generate new data.

    ----David Foster, Generative deep learning, 2019

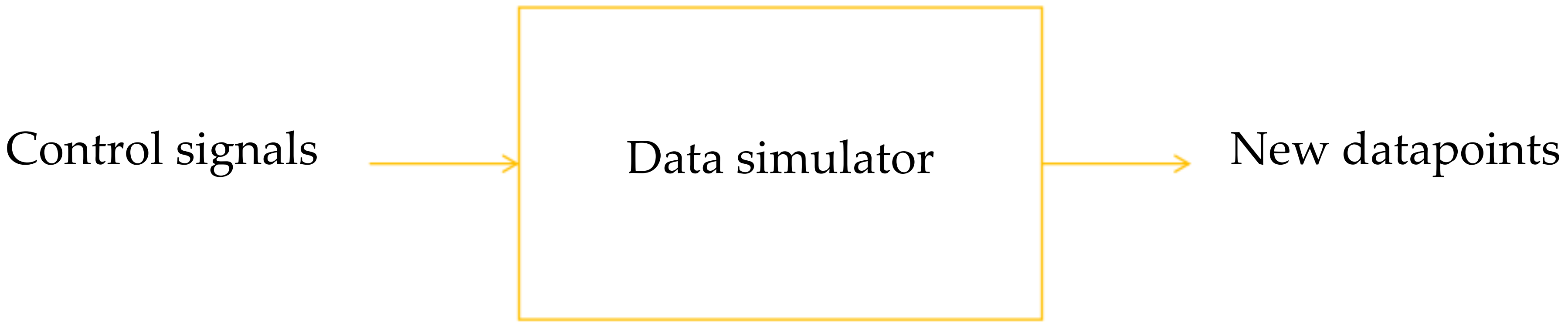


Data Generation "Simulator"

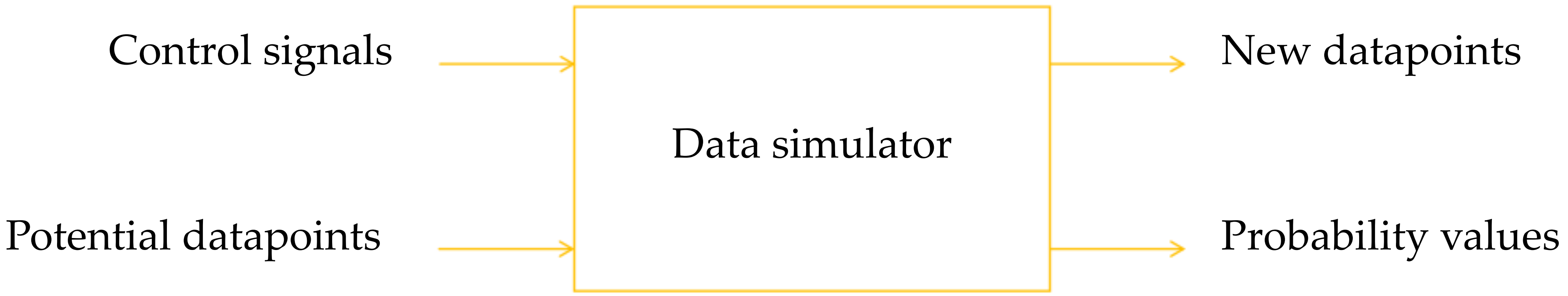
- Rethink from an application view…



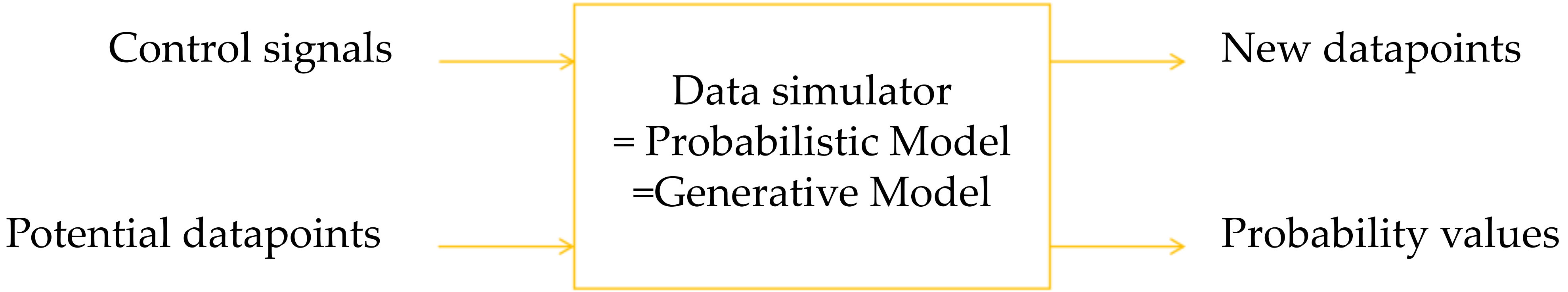
Data Generation "Simulator"



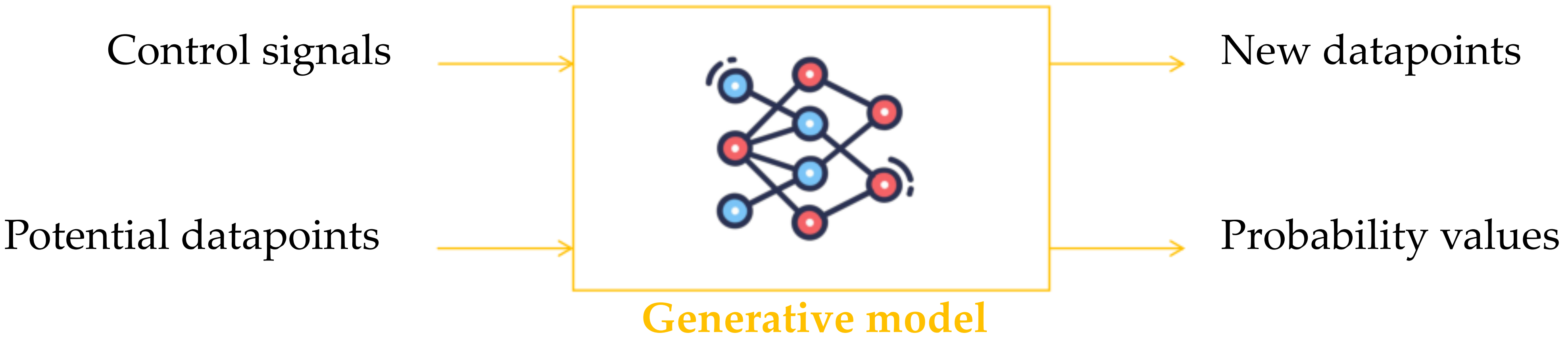
Data Generation "Simulator"



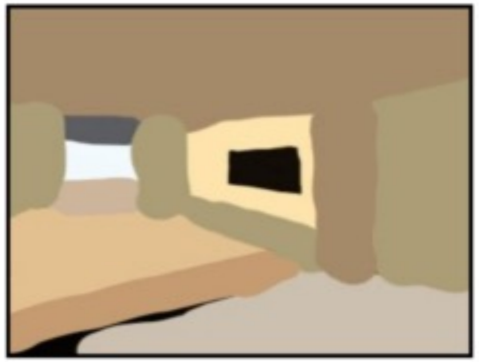
Data Generation “Simulator”



Data Generation "Simulator"



Apply to Real-World Data Generation



Generative model
of realistic images

Generate

Stroke paintings to realistic images
Meng, et al., 2021

"Ace of Pentacles"

Generative model
of paintings

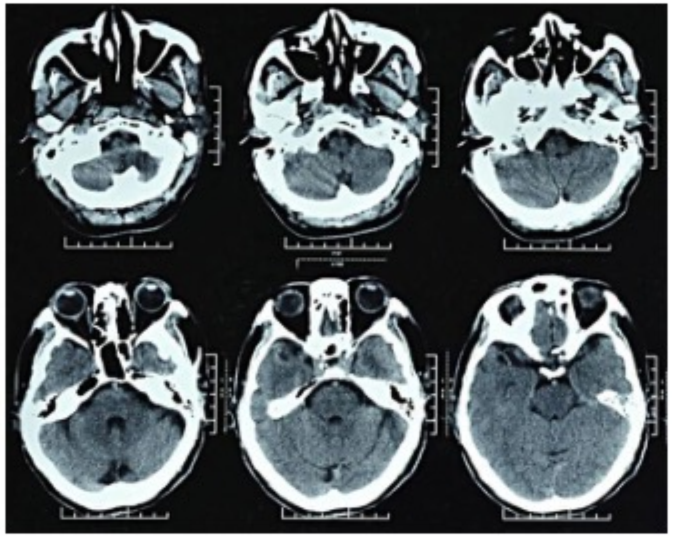
Generate

Language-guided artwork creation
RiversHaveWings, https://chainbreakers.kath.io

Apply to Solve Inverse Problems

Generative model
of medical images

Generate



Medical image reconstruction
Song et al., ICLR 2022

Apply for Outlier Detection

High probability

Generative model
of traffic signs

Low probability

Outlier detection
Song et al., ICLR 2018

Discriminative vs. Generative

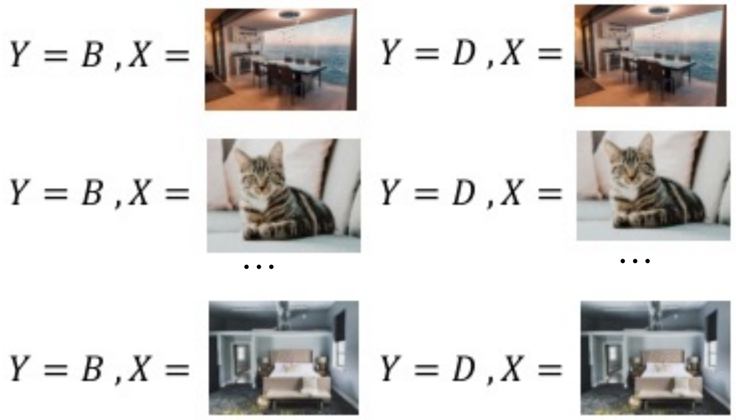
- **Discriminative** : classify bedroom vs. dining room

- The input image X is given
- **Goal** : a good decision boundary, via **conditional distribution** over label Y

$P(Y = \text{Bedroom} \mid X =$  $) = 0.0001$

- E.g., SVM, logistic regression, CNN, etc.

- **Generative** : generate X



- The input X is not given
- Generative models target **p(x), p(x,y), or p(x|c).**

$P(Y = \text{Bedroom}, X =$  $) = 0.0002$

- E.g., naïve Bayes, hidden Markov, etc.

Discriminative vs. Generative

- Joint and conditional are both related via **Bayes Rule**:

$P(Y = \text{Bedroom} \mid X =$  $) =$

Numerator: $P(Y = \text{Bedroom}, X =$  $)$

Denominator: $P(X =$  $)$

- **Discriminative**
  - Y is simple; X is always given, so not need to model $P(X =$  $)$

  - However, it cannot well handle missing data $P(Y = \text{Bedroom} \mid X =$  $)$

Discriminative vs. Generative

- **Generative Models**
  - Class conditional is possible:

$P(X =$  $\mid Y = \text{Bedroom}) =$ $P(Y = \text{Bedroom}, X =$  $)$ / $P(Y = \text{Bedroom})$

  - In the real world, it's often useful to condition on rich side information Y

$P(X =$  $\mid Y = $ "A black table with 6 chairs"$)$

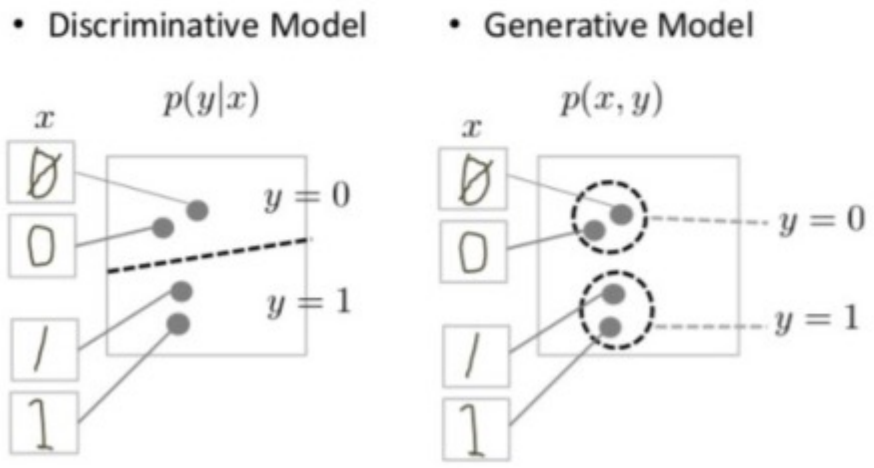
- p(y | x) is conditional modeling, but it is conventionally called discriminative for prediction tasks.

$\boxed{P(Y = \text{Bedroom}} \mid X =$  $)$

Recall
The probabilistic distribution

Discriminative vs. Generative

- Formally, given a set of data instances X and a set of labels Y:
  - Generative models capture the joint probability $p(X,Y)$, or just $p(X)$ if there are no labels
    - Modeling the distribution of individual classes
  - Discriminative models capture the conditional probability $p(Y|X)$
    - Modeling the boundaries between classes



Discriminative and generative models of handwritten digits

Discriminative vs. Generative

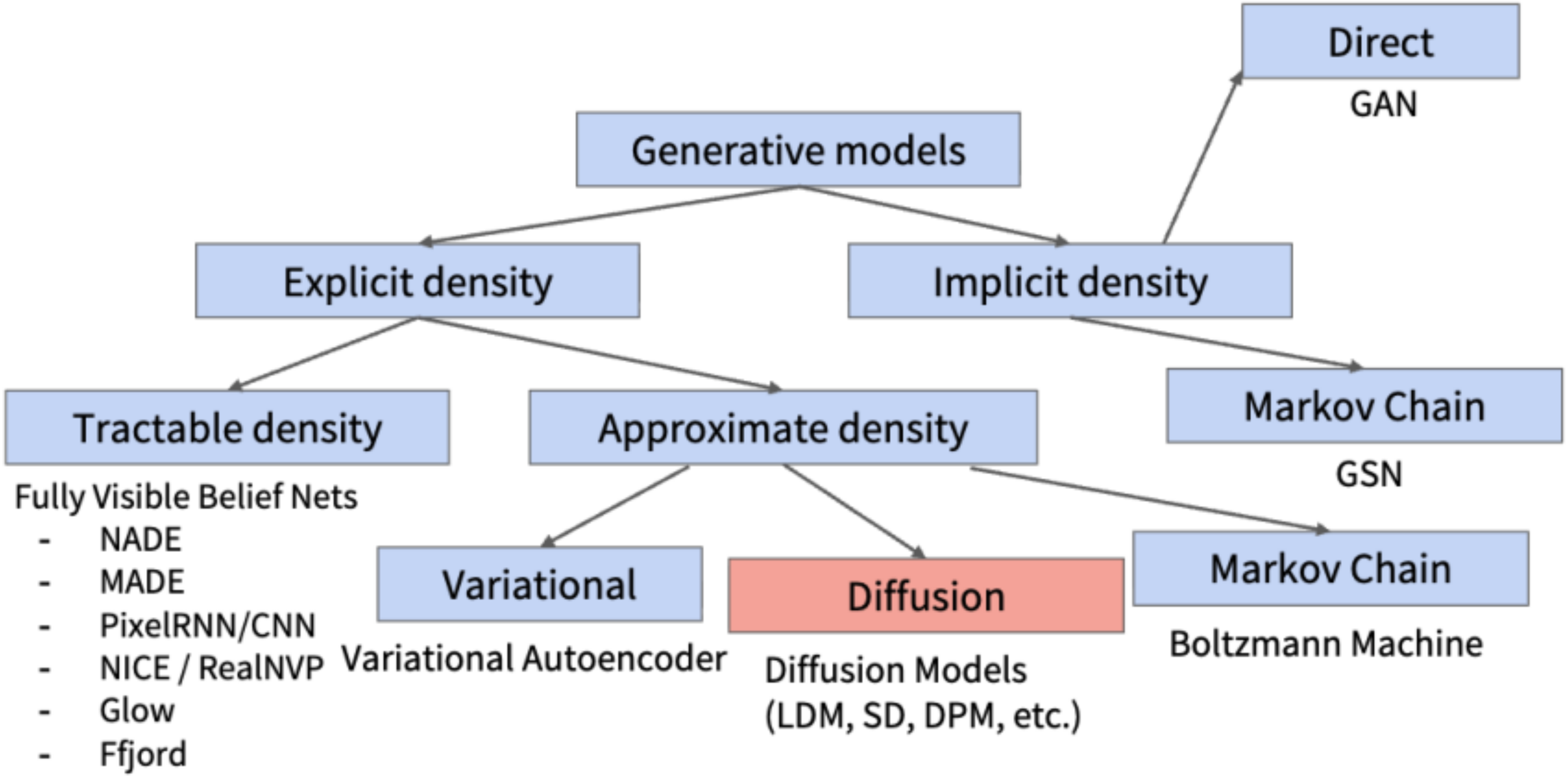
Typical tendencies under matched capacity/task

- **Discriminative**
  - Pros:
    - Simple and direct
    - Require fewer samples, lower training cost
    - High performance on classification tasks
  - Cons:
    - Capture limited information about data distribution
    - Cannot directly perform data generation or outlier detection

- **Generative**
  - Pros:
    - Capture richer information in data
    - More versatile: can handle generation, unsupervised learning, anomaly detection
    - Can also be used for discriminative tasks
  - Cons:
    - Require more samples and higher computational cost
    - Hallucination
    - Training is often more complex

Taxonomy of Generative Models

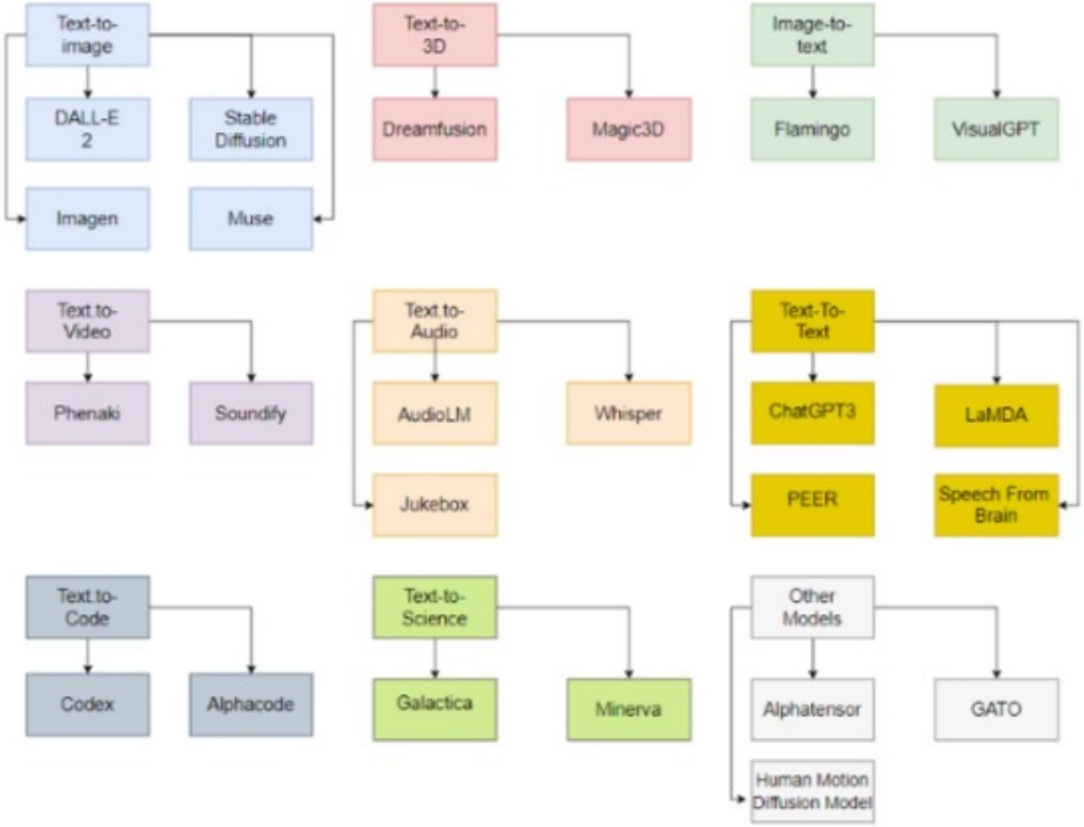
- From a **probabilistic modeling perspective**



Tutorial on Generative Adversarial Networks
Ian Goodfellow, 2017

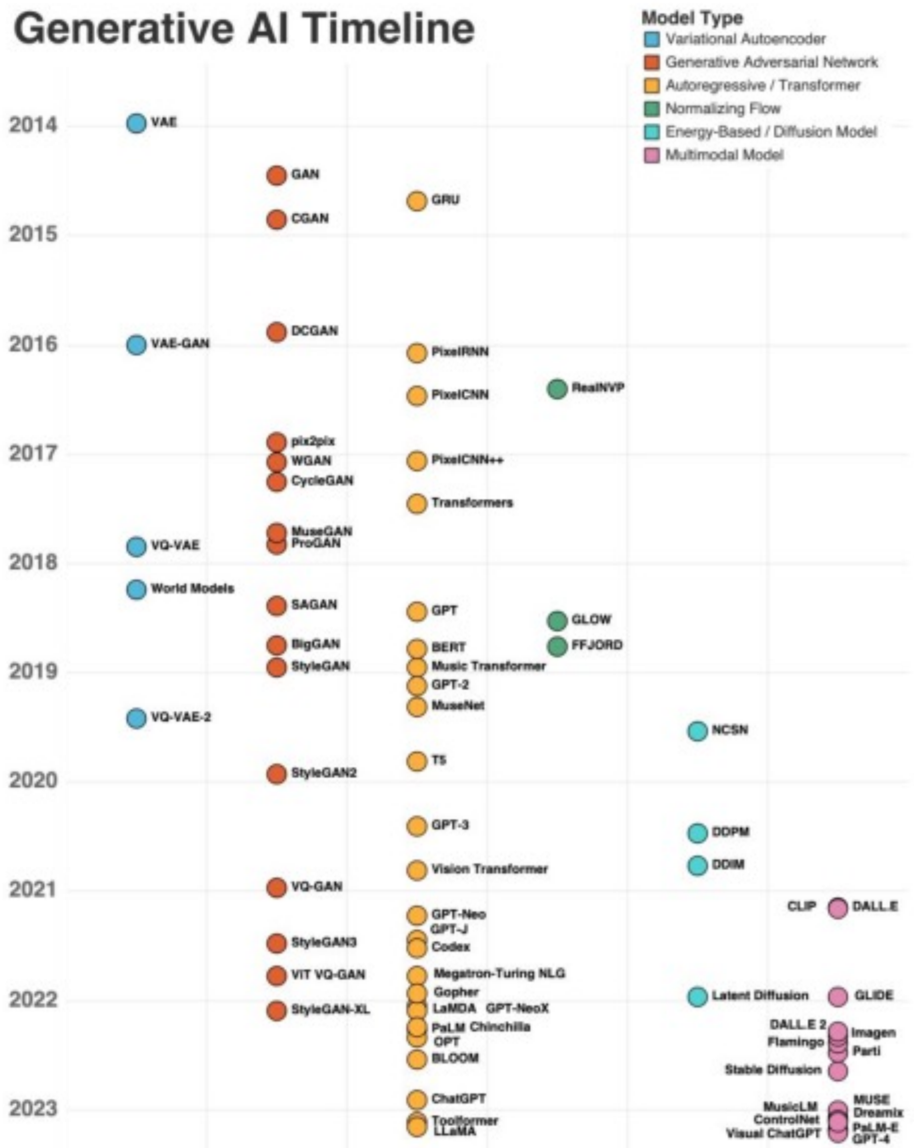
Taxonomy of Generative Models

- From an input/generated data modality perspective



Generative Models: A Timeline

- Recent models by date of release



Allena Venkata Sai Abhishek, 2023

Properties of Generative Models

- Common criteria
  - **Stability**
    - How easily can the model be trained?
  - **Sample quality & diversity**
    - Are the generated samples realistic and varied?
  - **Efficiency**
    - Is the model computationally efficient in training and inference?
  - **Likelihood estimation**
    - Can the model assign probabilities to samples?
  - **Capacity**
    - Can the model represent and generate complex data?
  - **Flexibility**
    - How constrained is the model by its structure or assumptions?
  - **Evaluability**
    - How easy is it to evaluate and compare this model?

	Stability	Quality	Diversity	Efficiency	Likelihood
VAE	☺	☹	☺	☺	☺
GAN	☹	☺	☹	☺	☹
DM	☺	☺	☺	☹	😐
Flow	😐	😐	☺	☺	☺

There is no one-for-all solution

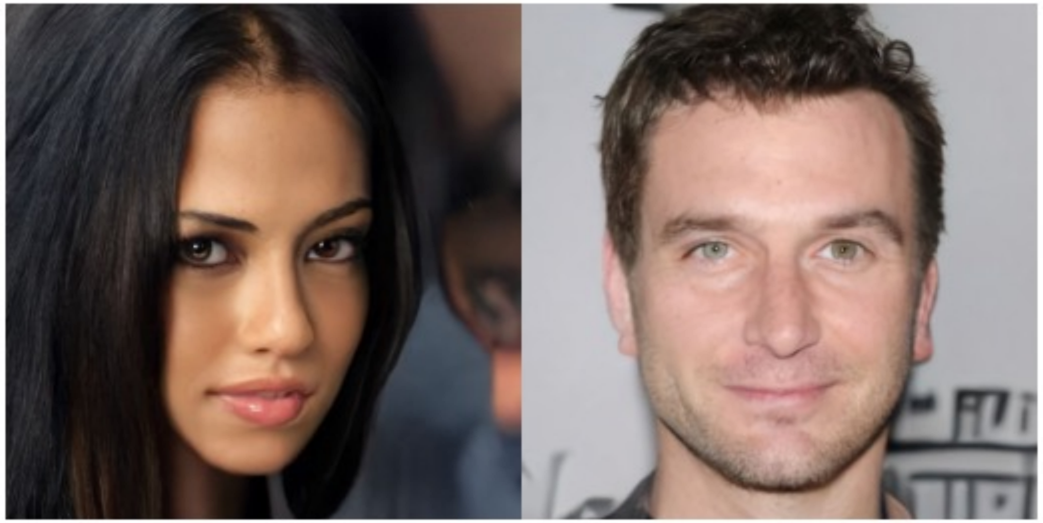
Recent Progress: Image Generation

- Face generation via VAE and GANs

Ian Goodfellow, 2019

Recent Progress: Image Generation

- Face generation via diffusion models



Score-Based Generative Modeling through Stochastic Differential Equations
Song et al., 2021

Flux. 1
2024

Recent Progress: Image Generation

- Face dataset generation

MorphFace
Mi et al., 2026

Recent Progress: Text-to-Image

- T2I User input:
  - *"An astronaut riding a horse"*

Dall-E 2
2022

Recent Progress: Text-to-Image

- T2I User input:
  - *“A perfect Italian meal”*

Stable Diffusion 3
2024

Recent Progress: Text-to-Image

- T2I User input:
  - “泰迪熊穿着戏服，站在太和殿前唱京剧”

文 言
2023

Recent Progress: Text-to-Image

- T2I User input:
  - *"A minimap diorama of a cafe adorned with indoor plants. Wooden beams crisscross above, and a cold brew station stands out with tiny bottles and glasses"*

Dall-E 3

(Built natively on ChatGPT)

Recent Progress: Text-to-Image

- T2I User input:
  - 根据文字描述虚构漫画：A story about LeCun, Hinton, and Bengio winning the Nobel Prize

Recent Progress: Inverse Problems

- Inpainting / Super-Resolution / Restoration

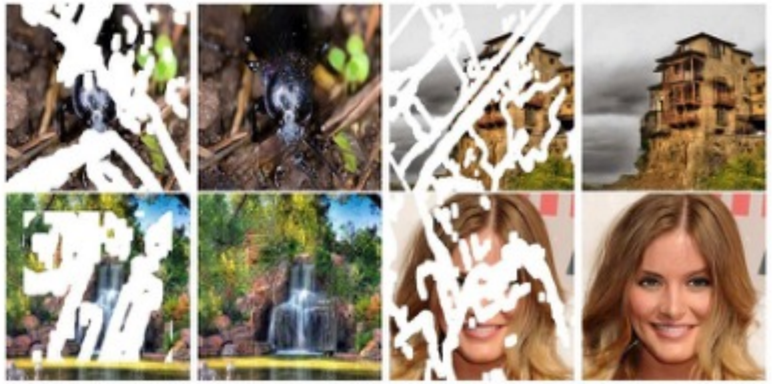
P(high resolution | low resolution)

Menon et al, 2020

P(color image | greyscale)

Antic et al., 2020

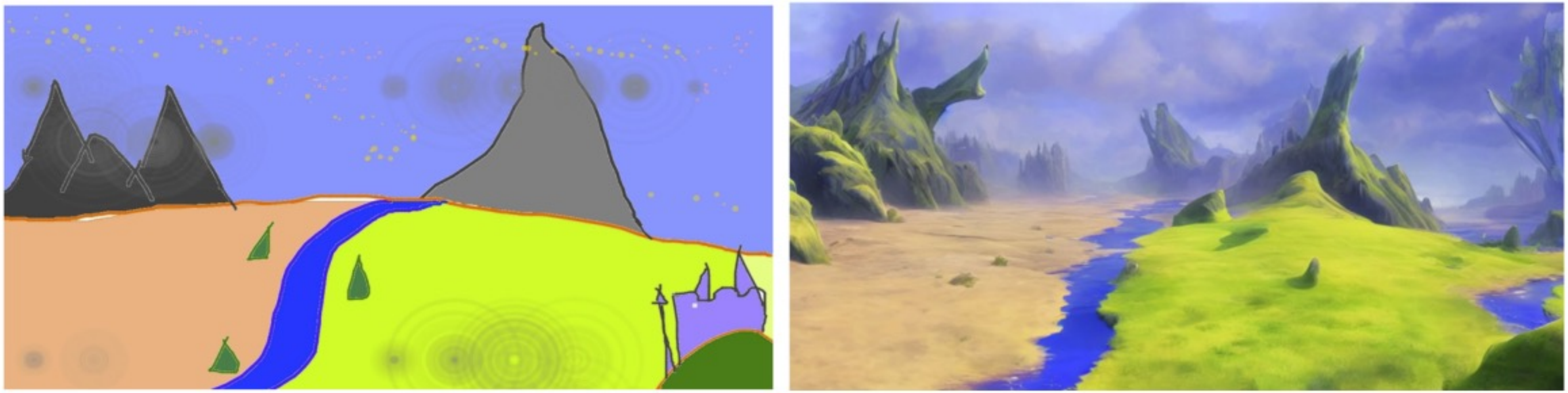
P(full image | mask)



Liu et al, 2018

Recent Progress: Inverse Problems

- Sketch-to-image



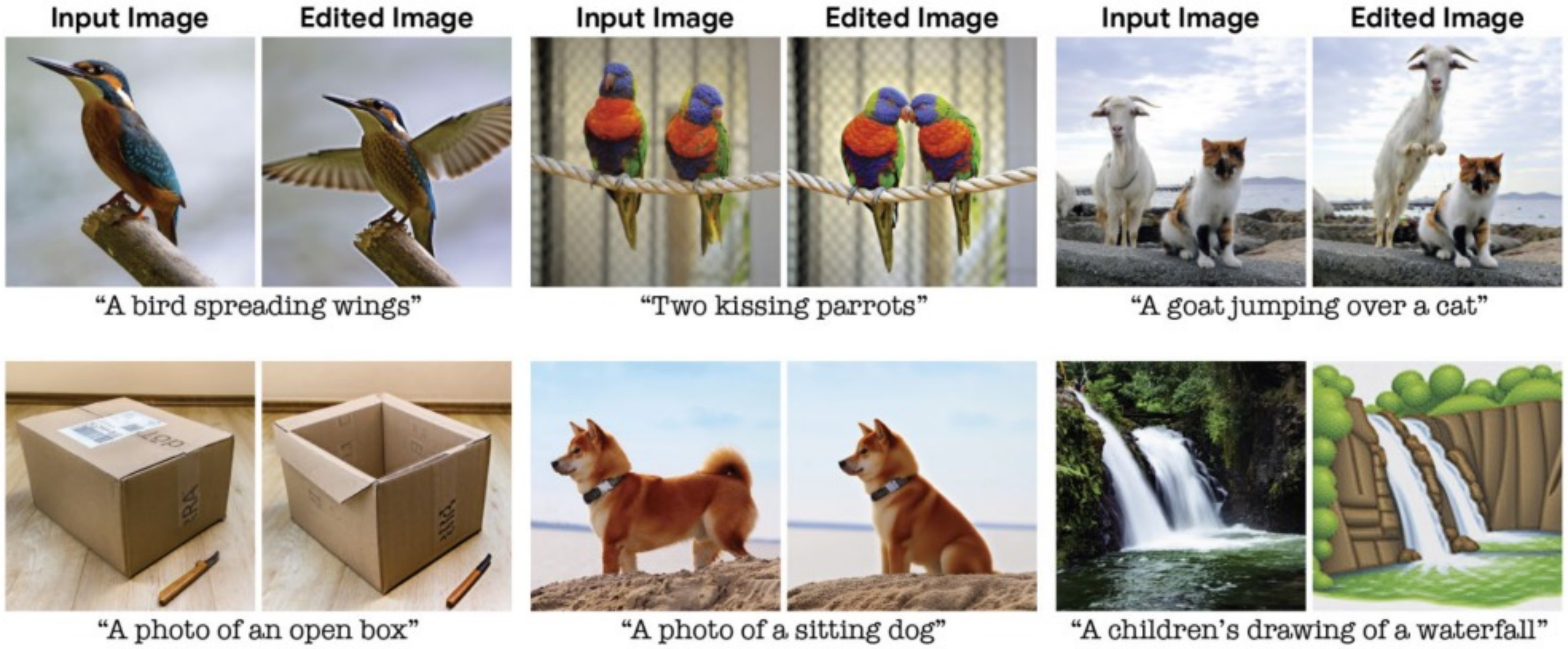
Recent Progress: Inverse Problems

- Sketch-to-image

SDEdit
Meng et al., 2021

Recent Progress: Inverse Problems

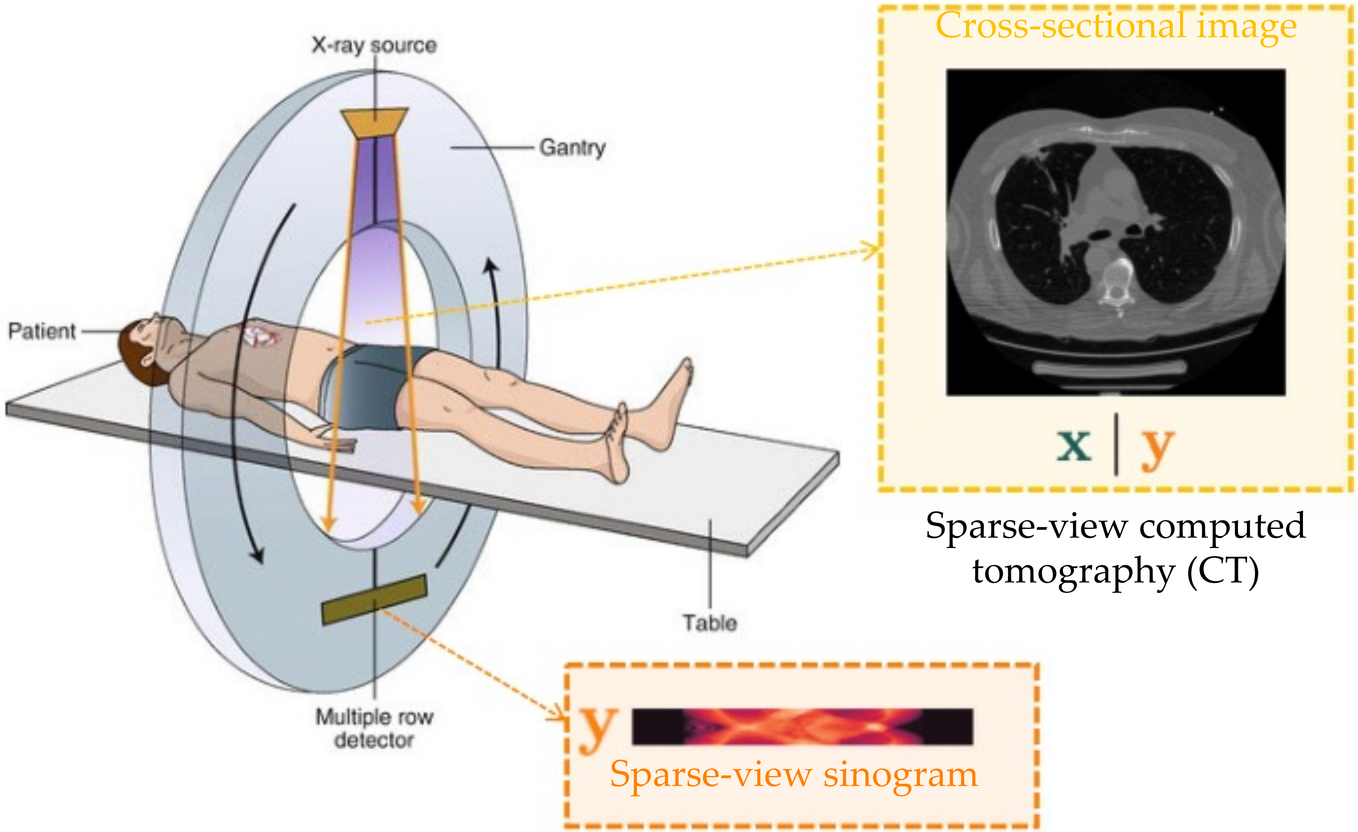
- Text-based image editing



Kawar et al., 2023

Recent Progress: Inverse Problems

- Medical image reconstruction

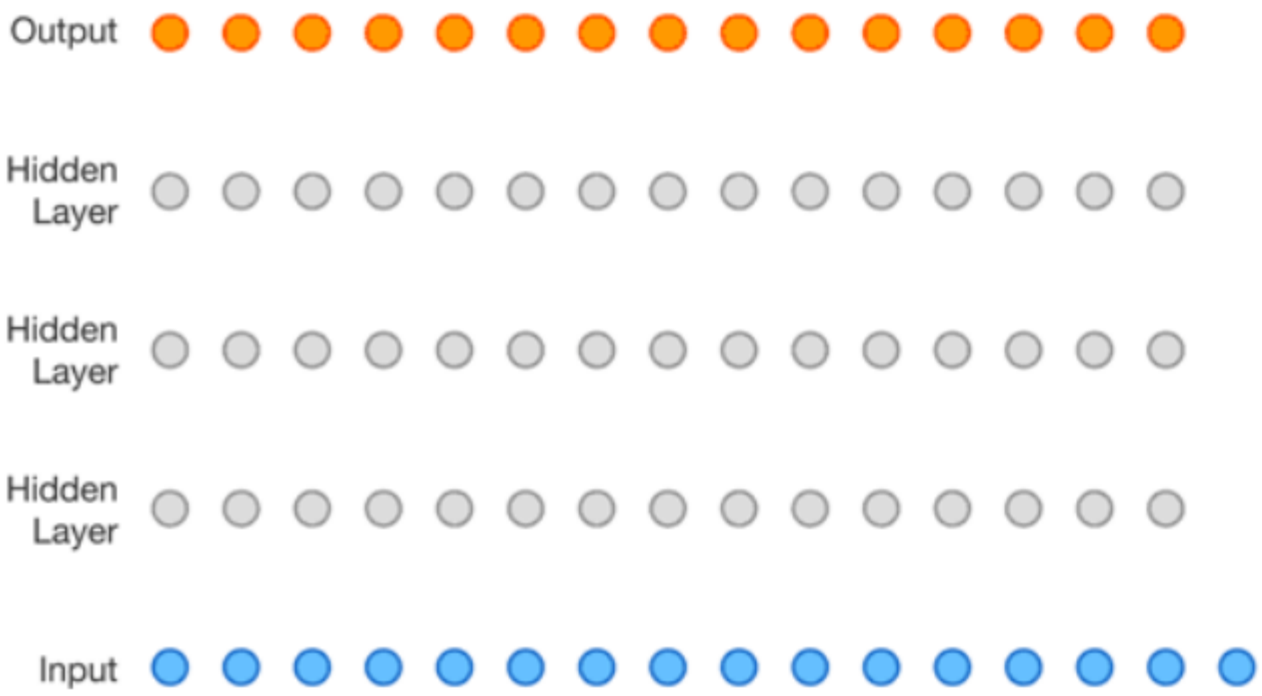


Sparse-view computed tomography (CT)

Forward model $p(\mathbf{y} \mid \mathbf{x})$ is given by physical simulation

Recent Progress: Audio & Speech

- Generative model of speech signals



WaveNet
van den Oord et al., 2016

Text to Speech

Parametric

Concatenative

WaveNet

Unconditional

Music

Recent Progress: Language Generation

- Text auto completion
  - Next-word prediction: $P(\text{next word} \mid \text{previous words})$

Custom prompt

To get an A+ in deep generative models, students have to

Completion

To get an A+ in deep generative models, students have to be willing to work with problems that are a whole lot more interesting than, say, the ones that most students work on in class. If you're a great student, the question above can be avoided and you'll be able to do great work, but if you're not, you will need to go beyond the basics before getting good.

Now to be clear, this advice is not just for the deep-learning crowd; it is good advice for any student who is taking his or her first course in machine learning.

The key point is that if you have a deep, deep brain of a computer scientist, that's just as important to you.

Radford et al., 2019

Recent Progress: Language Generation

- GPT

How to learn generative models well?

Learning generative models well is a big journey, but you can approach it step by step depending on your background and goals. Here's a structured roadmap I'd suggest:

---

1. Build Core Foundations

Before diving into generative models, strengthen the basics:

- **Probability & Statistics**: distributions, Bayes' rule, KL divergence, entropy.
- **Linear Algebra & Optimization**: matrix operations, eigen-decomposition, gradient descent, convexity.
- **Deep Learning Fundamentals**: backpropagation, CNNs, RNNs, transformers.

👉 Recommended: *Deep Learning* by Goodfellow, Bengio, Courville (especially Chapters 3–5, 17).

---

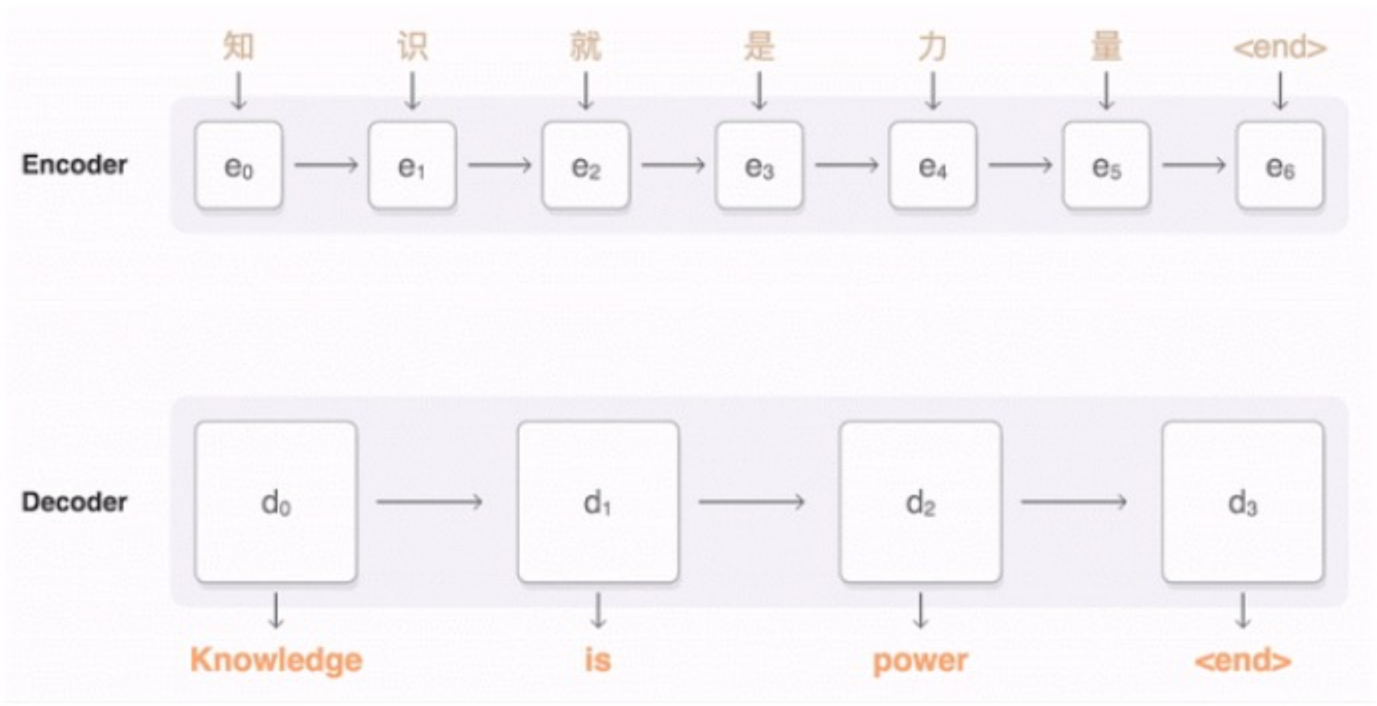
2. Understand the Landscape of Generative Models

Generative models aim to learn $p(x)$ or $p(x|y)$. The major families are:

- **Explicit likelihood models**
  - Gaussian Mixture Models (classical start)
  - Variational Autoencoders (VAE)
  - Normalizing Flows (e.g., RealNVP, Glow)
- **Implicit models**
  - Generative Adversarial Networks (GANs, StyleGAN, BigGAN)
  - Diffusion Models (DDPM, Stable Diffusion, Imagen)

Recent Progress: Language Generation

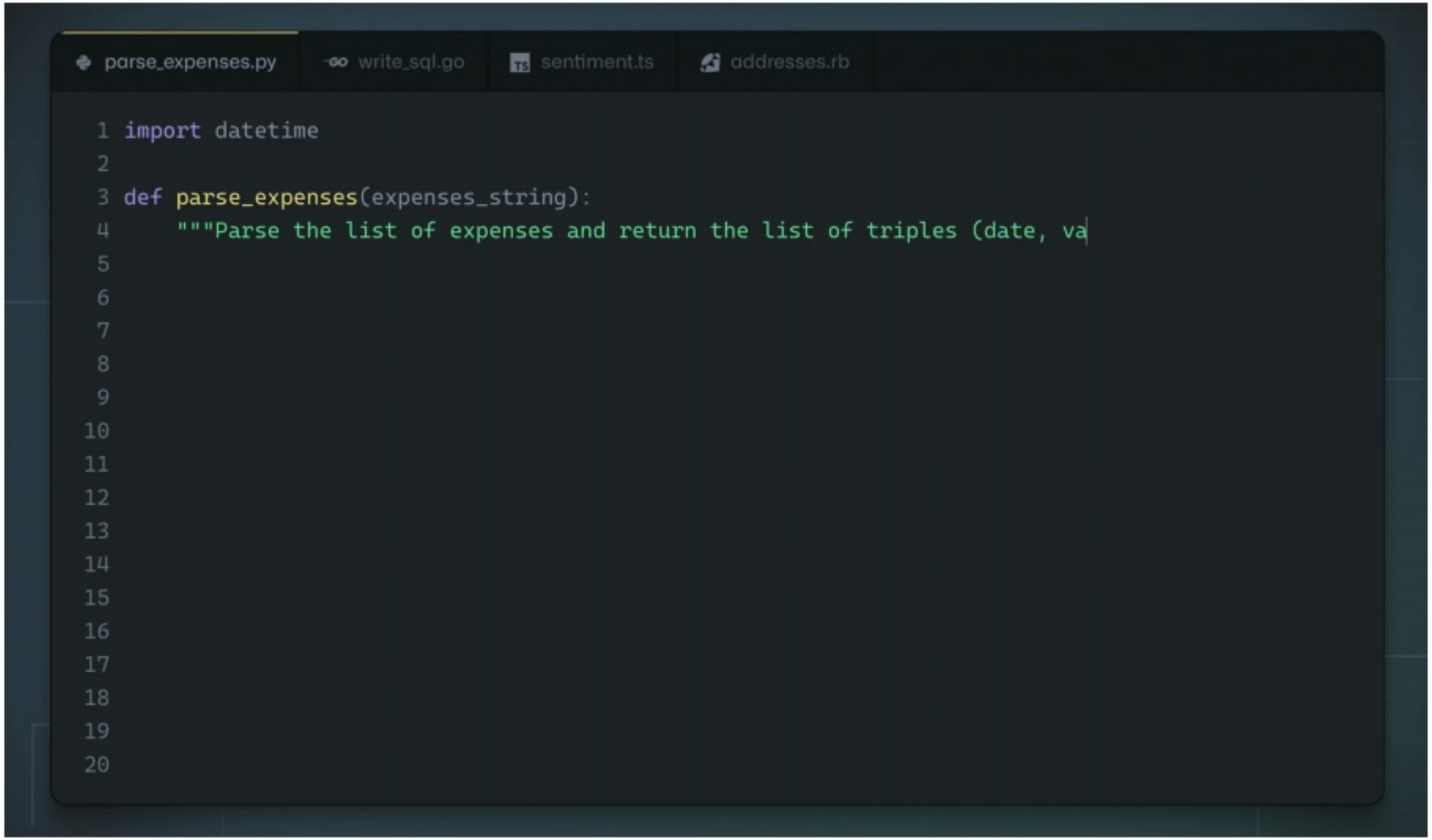
- Machine translation
  - Conditional generative model $P(\text{English text}|\ \text{Chinese text})$



Google AI research

Recent Progress: Code Generation

- Github Copilot
- OpenAI Codex



Recent Progress: Video Generation

- T2V User Input:
  - “Suddenly, the walls of the embankment broke and there was a huge flood.”

Pika
Pika Labs

Recent Progress: Video Generation

- T2V User Input:
  - "A couple sledding down a snowy hill on a tire roman chariot style."

Pika
Pika Labs

Recent Progress: Video Generation

- Video and audio generation

Recent Progress: Video Generation

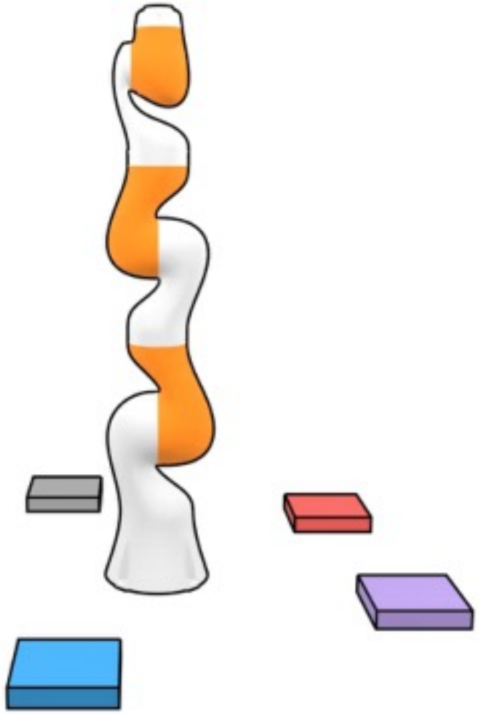
- Talking head

Zhong et al., 2026

Recent Progress: Robotics

- Imitation Learning
  - Conditional generative model $P(\text{actions} \mid \text{past observations})$

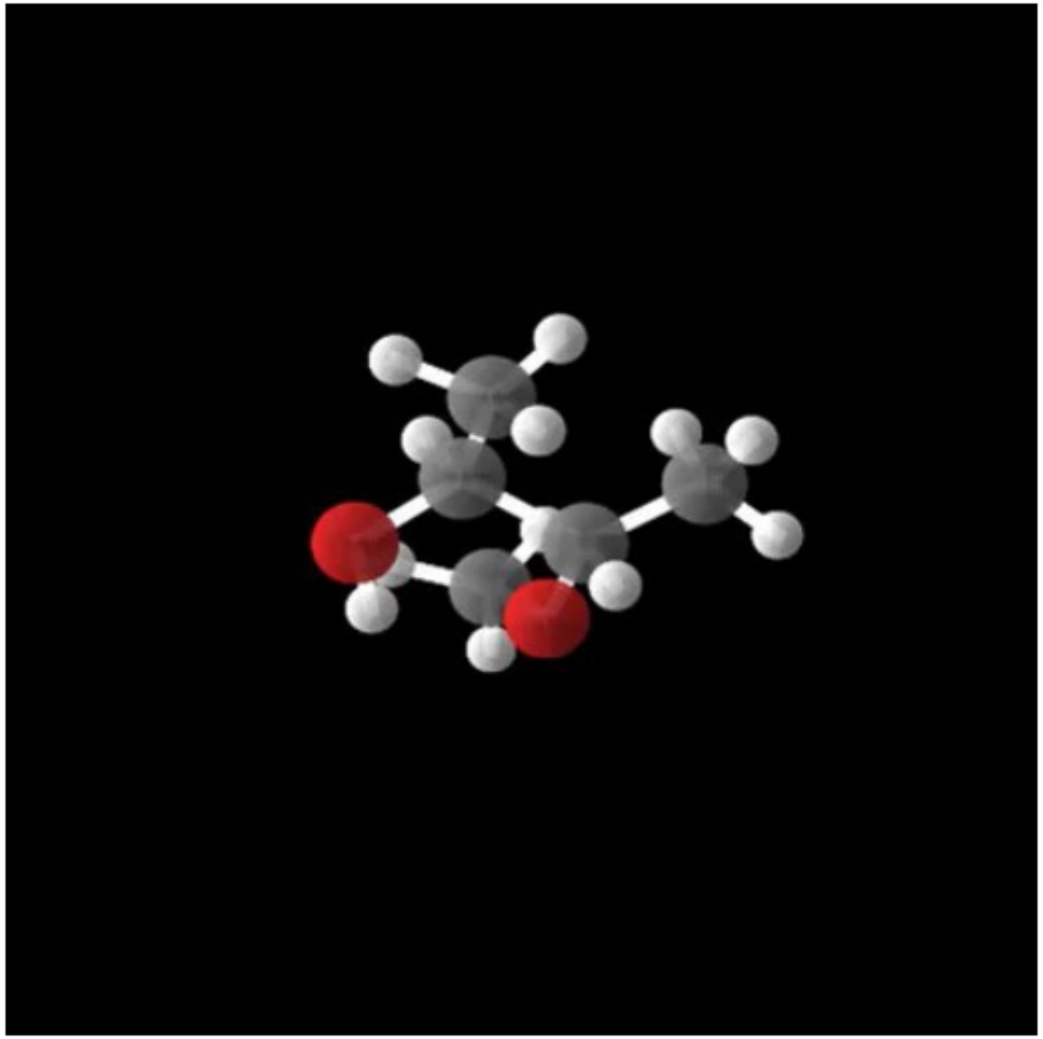
Li et al., 2017



Janner et al., 2022

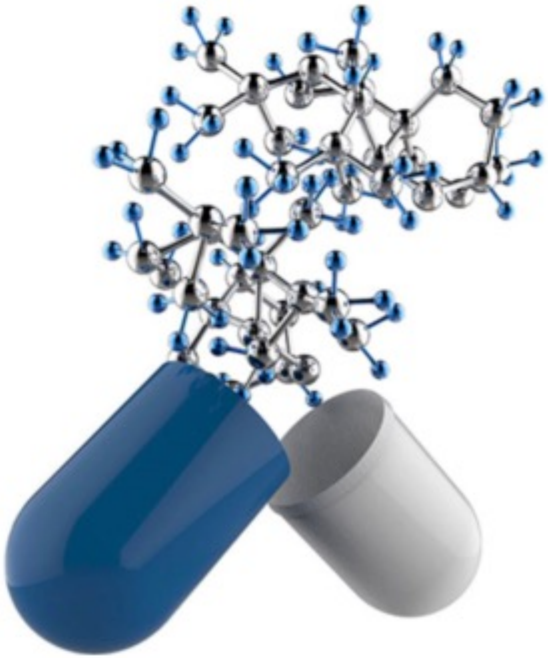
Recent Progress: AI4S

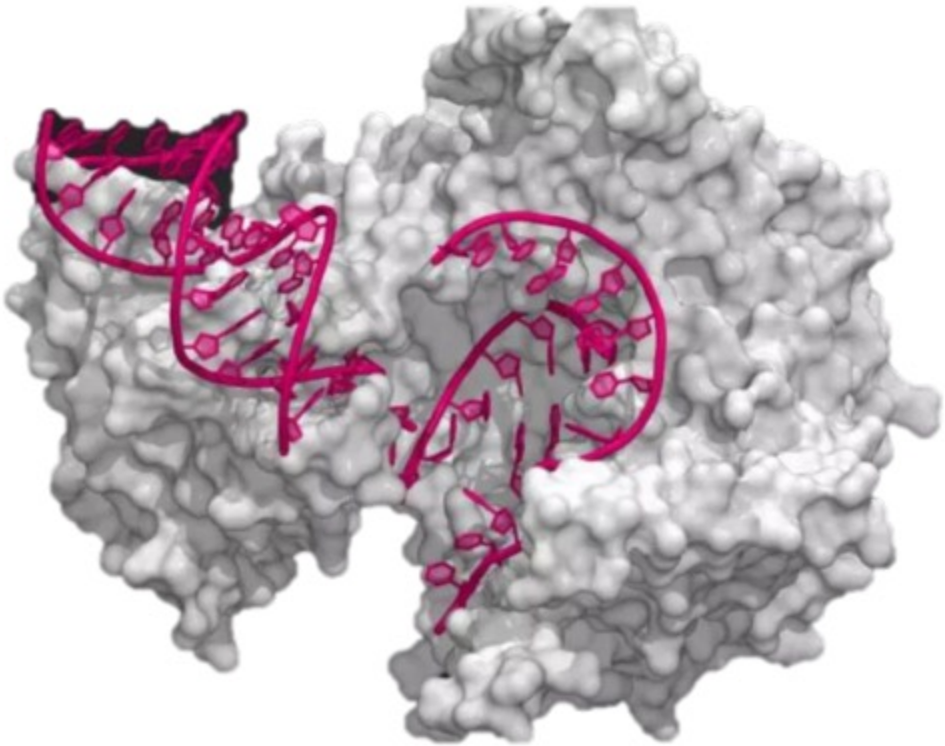
- Molecule generation



Recent Progress: AI4S

- Bioinformatics

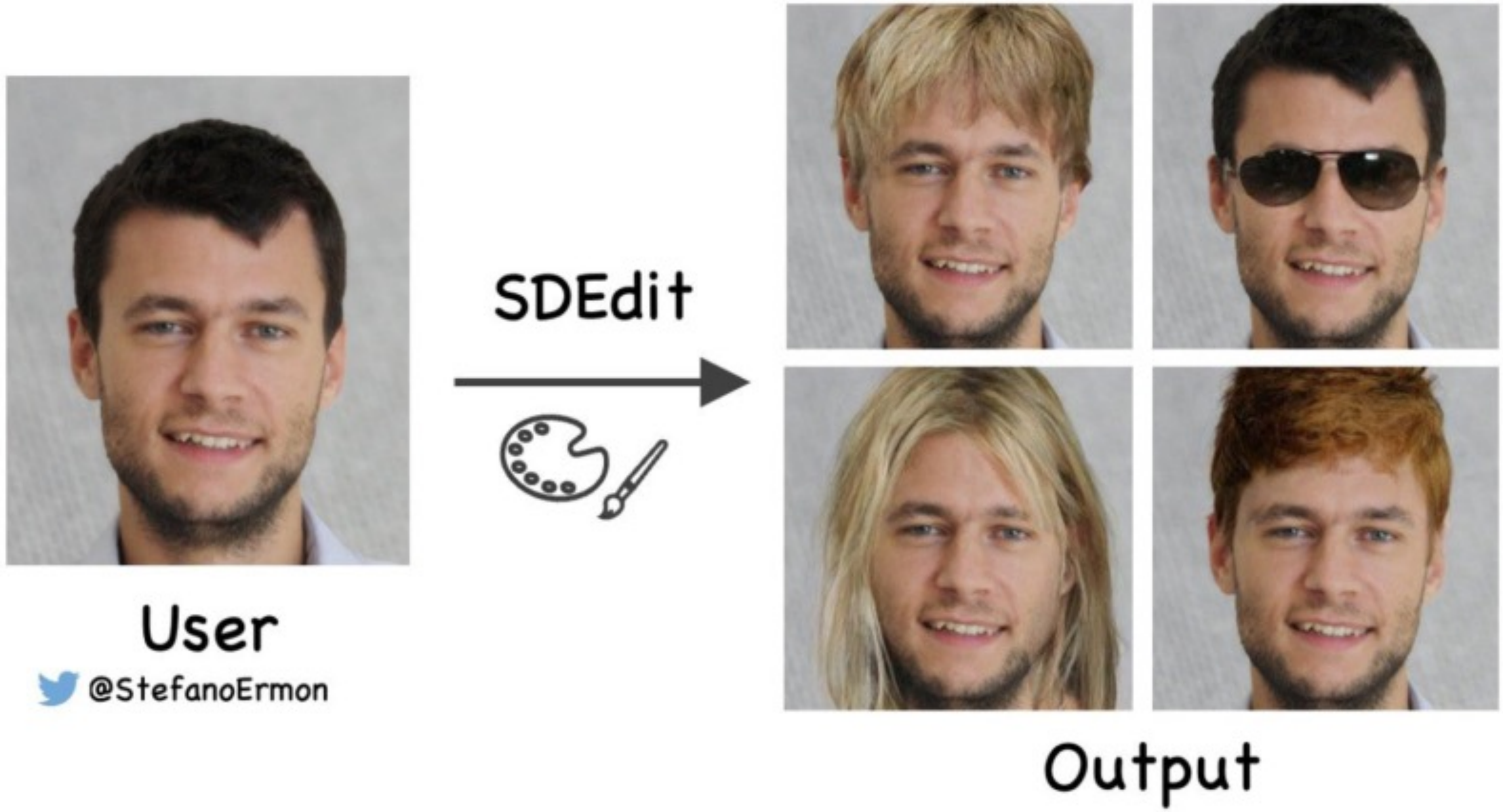




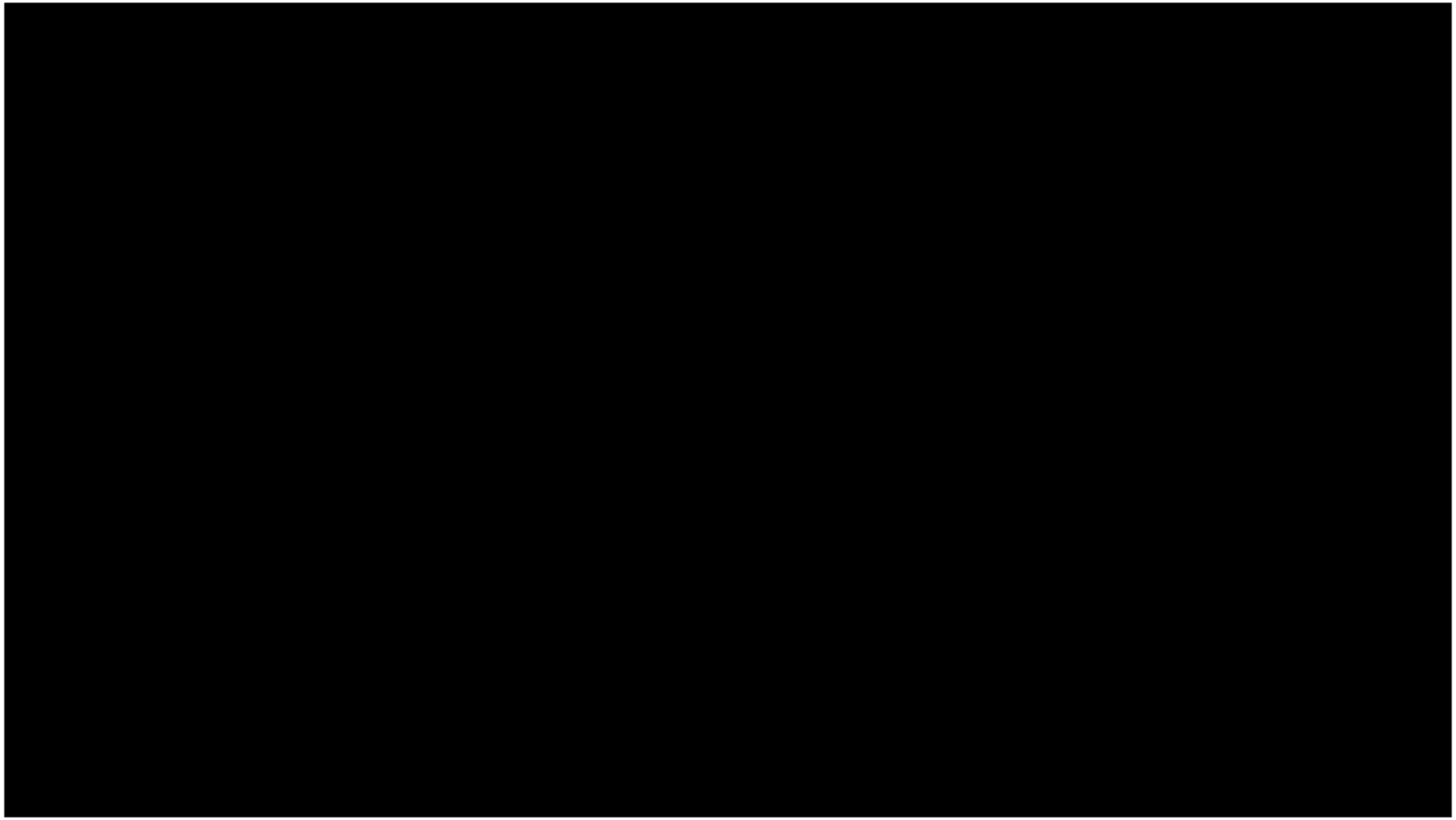
ProteinGAN
Repecka et al., Nature Machine Intelligence, 2021

Recent Progress: DeepFake

- Which image is real?



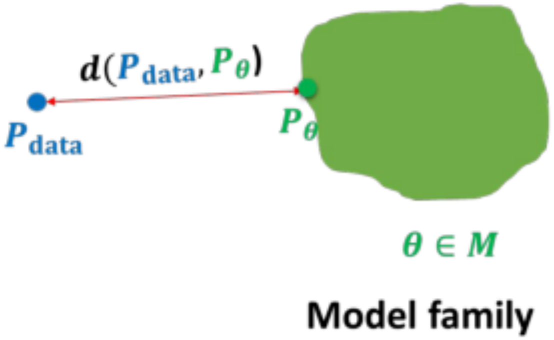
Recent Progress: DeepFake



Roadmap and Key Challenges

- **Representation**
  - How do we model the joint distribution of many random variables?
  - Need compact representation
- **Learning**
  - What is the right way to compare probability distributions?

$\mathbf{x}_i \sim P_{\text{data}}$
$i = 1, 2, \dots, n$



- **Inference**
  - How do we invert the generation process (e.g., vision as inverse graphics)?
  - Unsupervised learning: recover high-level descriptions (features) from raw data

Syllabus

- Basic concepts of generative models
- Generative models in statistical learning
  - Naive Bayesian models, Hidden Markov models, Gaussian mixture models
  - Probabilistic latent semantic analysis models (PLSA), latent Dirichlet allocation
- **Generative models in deep learning**
  - Autoregressive models
  - Variational auto-encoders (VAEs)
  - Generative adversarial networks (GANs)
  - Flow-based models
  - Energy-based models (EBMs)
  - Score-based / diffusion generative models (DMs)
  - Large-language models (LLMs)
- Applications in different areas
  - Image processing: image generation
  - NLP: text generation
  - Speech recognition: audio generation
  - Video processing: video generation
  - Drug design & discovery: molecule generation

Syllabus

- Basic concepts of generative models
- Generative models in statistical learning
  - Naive Bayesian models, Hidden Markov models, Gaussian mixture models
  - Probabilistic latent semantic analysis models (PLSA), latent Dirichlet allocation
- **Generative models in deep learning**
  - Autoregressive models *~2011*
  - Variational auto-encoders (VAEs) *~2013*
  - Generative adversarial networks (GANs) *~2014*
  - Flow-based models *~2015*
  - Energy-based models (EBMs) *~2016*
  - Score-based / diffusion generative models (DMs) *~2019*
  - Large-language models (LLMs) *~2020*

Approximate period to prominence in deep-learning era

- Applications in different areas
  - Image processing: image generation
  - NLP: text generation
  - Speech recognition: audio generation
  - Video processing: video generation
  - Drug design & discovery: molecule generation

Syllabus

- Basic concepts of generative models
- Generative models in statistical learning
  - Naive Bayesian models, Hidden Markov models, Gaussian mixture models
  - Probabilistic latent semantic analysis models (PLSA), latent Dirichlet allocation
- **Generative models in deep learning**
  - Autoregressive models
  - Flow-based models
  - Variational auto-encoders (VAEs)
  - Generative adversarial networks (GANs)
  - Energy-based models (EBMs)
  - Score-based / diffusion generative models (DMs)
  - Large-language models (LLMs)
- Applications in different areas
  - Image processing: image generation
  - NLP: text generation
  - Speech recognition: audio generation
  - Video processing: video generation
  - Drug design & discovery: molecule generation

Explicit likelihood-based models: $p_\theta(x)$

– *Latent variable models:* $p_\theta(x,z) = p_\theta(x|z)p(z)$, $p_\theta(x) = \int p_\theta(x \mid z)p(z)dz$

– *Implicit models: without* $p(x)$

Energy-based & Extensions: $E_\theta(x)$ *or* $\nabla_x \log p(x)$

Prerequisites

- **Mathematics**
  - Basic knowledge of probability and calculus
    - Random variables, independence, conditional independence
    - Gradients, gradient-descent optimization
    - Bayes rule, chain rule, change of variables formulas
- **Computer science**
  - basic data structures/algorithms
  - basic knowledge of machine learning/deep learning
- **Programming language, preferably Python**

Course Goal

- **Understand**
  - Grasp the fundamentals of generative models
  - Read the latest research papers without obstacles
- **Familiarize**
  - Learn major generative models, especially deep ones, with their pros & cons
- **Recognize**
  - Identify the latest achievements and development trends in typical application areas
- **Explore**
  - If possible, try to conduct some research
  - Say, improve existing models or apply them to solve real problems

Grading Policies

- **Attendance: 20%**
- **Paper presentation: 40%**
  - Introduce at least one paper from top venues
- **Course article: 40%**
  - A survey or a research-style article related to generative models
  - Preferably use LaTeX

Logistics

Instructors

周水庚

D4021, No. 2 Interdisciplinary Building

Email: sgzhou@fudan.edu.cn

Homepage: http://admis.fudan.edu.cn/sgzhou

刘文辉

D4023, No. 2 Interdisciplinary Building

Email: wh_liu@fudan.edu.cn

The best way to reach us is email!

Logistics

- **Time**
  - 9:55 AM – 12:30 PM
    Fridays from September 11 to December 25, 2026
  - W1-W16
- **Venue**
  - JB 106

Recommended Texts

- *S. Bond-Taylor, A. Leach, Y. Long, et al. Deep Generative Modelling: A Comparative Review of VAEs, GANs, Normalizing Flows, Energy-Based and Autoregressive Models. arXiv, 2021.*
- *E. Nalisnick, A. Matsukawa, Y. W. Teh, et al. Do Deep Generative Models Know What They Don't Know? ICLR, 2019.*
- *R. Salakhutdinov. Learning deep generative models. Annual Review of Statistics and Its Application, 2015.*
- *Kevin P. Murphy. Machine Learning: A Probabilistic Perspective. The MIT Press, 2012.*
- *David Foster. Generative Deep Learning. O'Reilly Media, 2019.*
- *I. Goodfellow, Y. Bengio, A. Courville (eds.), Deep Learning. The MIT Press, 2016.*
- *李航. 统计学习方法. 清华大学出版社, 2012.*

Recommended Top Venues

- **Journals**
  - Nature, Science, Nature Machine Intelligence
  - TPAMI, IJCV, AI, JMLR
  - TIP
- **Conferences**
  - NeurIPS, ICLR, ICML
  - ICCV, CVPR, ECCV
  - IJCAI, AAAI, MM, COLT, ACL, COLM

Thank you for listening!

Questions?