Introduction to Deep Generative Modeling

Lecture #1

HY-673 – Computer Science Dept, University of Crete

Professors: Yannis Pantazis & Yannis Stylianou

TAs: Michail Raptakis & Michail Spanakis

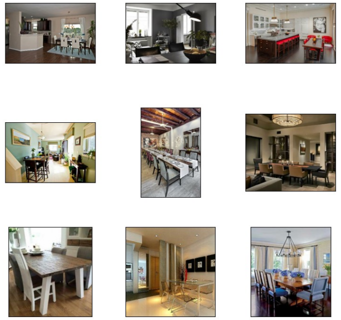
What is this course about?

✓ Statistical Generative Models

- ✓ A Generative Model (GM) is defined as a **probability distribution**, $\boldsymbol{p}(\boldsymbol{x})$.
  - ✓ A statistical GM is a *trainable probabilistic* model, $\boldsymbol{p}_{\boldsymbol{\theta}}(\boldsymbol{x})$.
  - ✓ A deep GM is a statistical generative model parametrized by a *neural network*.
  - ✓ $\boldsymbol{p}(\boldsymbol{x})$ and in many cases $\boldsymbol{p}_{\boldsymbol{\theta}}(\boldsymbol{x})$ are not analytically known. Only samples are available!
- ✓ **Data** ($\boldsymbol{x}$)**:** complex, (un)structured samples (e.g., images, speech, molecules, text, etc.)
- ✓ **Prior knowledge:** parametric form (e.g., Gaussian, mixture, softmax), loss function (e.g., maximum likelihood, divergence), optimization algorithm, invariance/equivariance, laws of physics, prior distribution, etc.

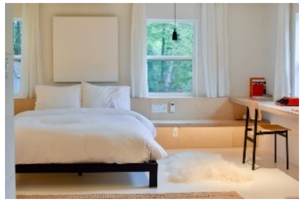
What is this course about?

- ✓ A dataset with images e.g., of bedrooms (LSUN dataset)



→ data distribution $p(x)$ or $p_{data}(x)$ or $p_d(x)$ → GM’s distribution $p_\theta(x)$ or $p_g(x)$

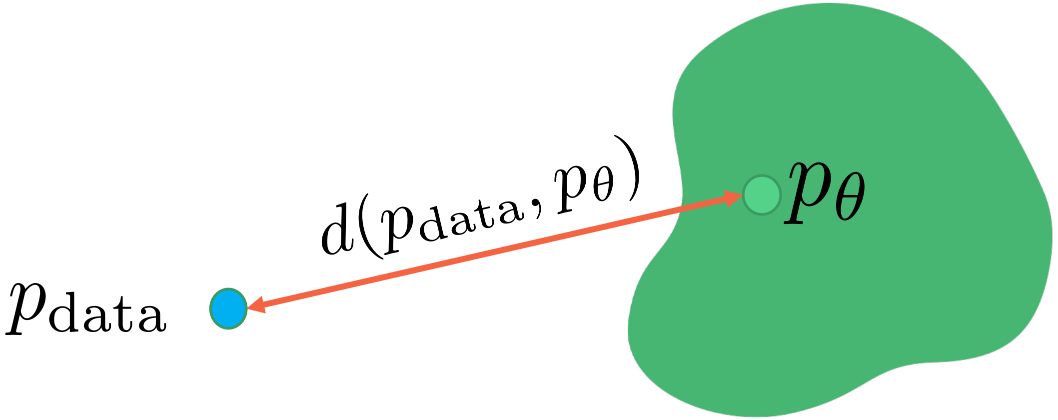
- ✓ ***<u>Goal:</u>*** Find $\theta \in \Theta$ such that $p_\theta(x) \approx p_d(x)$.
- ✓ It is *<u>generative</u>* because *sampling from $p_\theta(x)$ generates* new *<u>unseen</u>* images.

 …  $\sim \quad p_\theta(x)$

What is this course about?

Lecture #1

$x_i \sim p_{\text{data}}$
$i = 1, 2, \ldots, n$



Parametric Model Space

We will stydy:

- ✓ Families of Generative Models
- ✓ Algorithms to train these GMs
- ✓ Network architectures
- ✓ Loss functions & distances between probability density functions

What is this course about?

✓Conditional Generative Models

- ✓A conditional GM is defined as a **conditional probability distribution**, $p(x|y)$.
- ✓y: conditioning variable(s) (e.g., label/class, text, captions, speaker id, style, rotation, thickness, …)

$\sim \quad p_\theta(x|y),\ y$: *digit label*

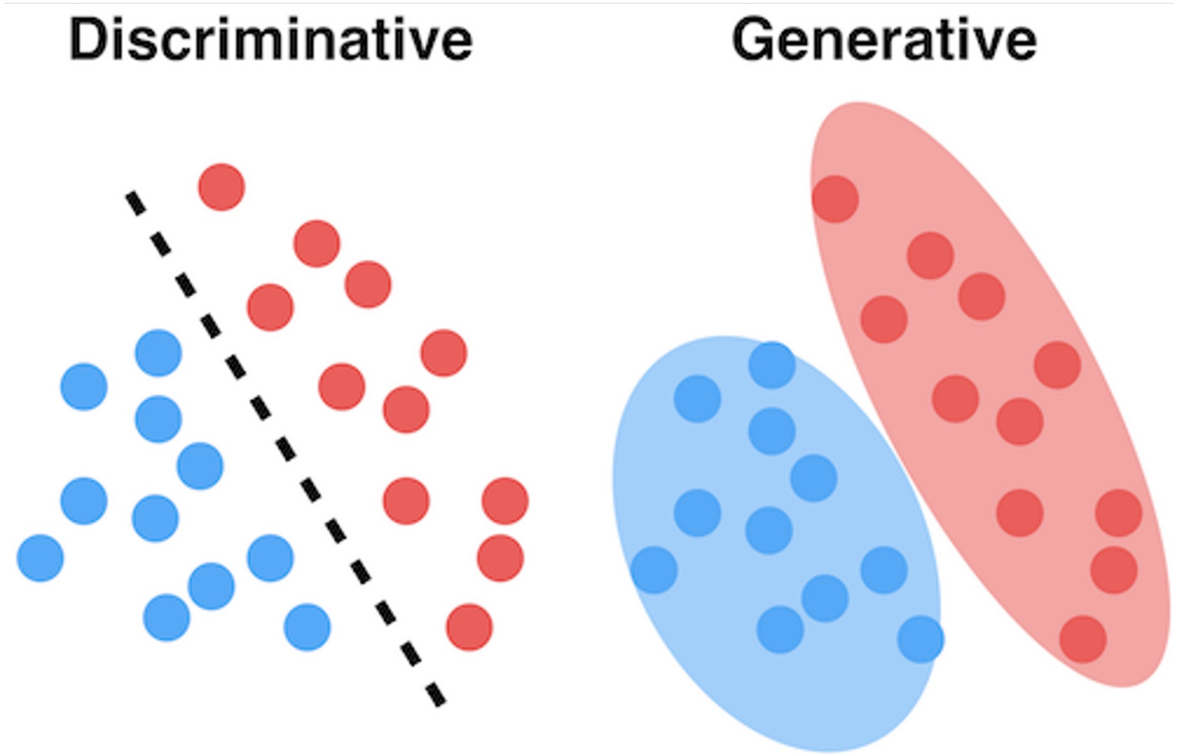
Discriminative vs Generative Models

Lecture #1

Data: x

Label: y

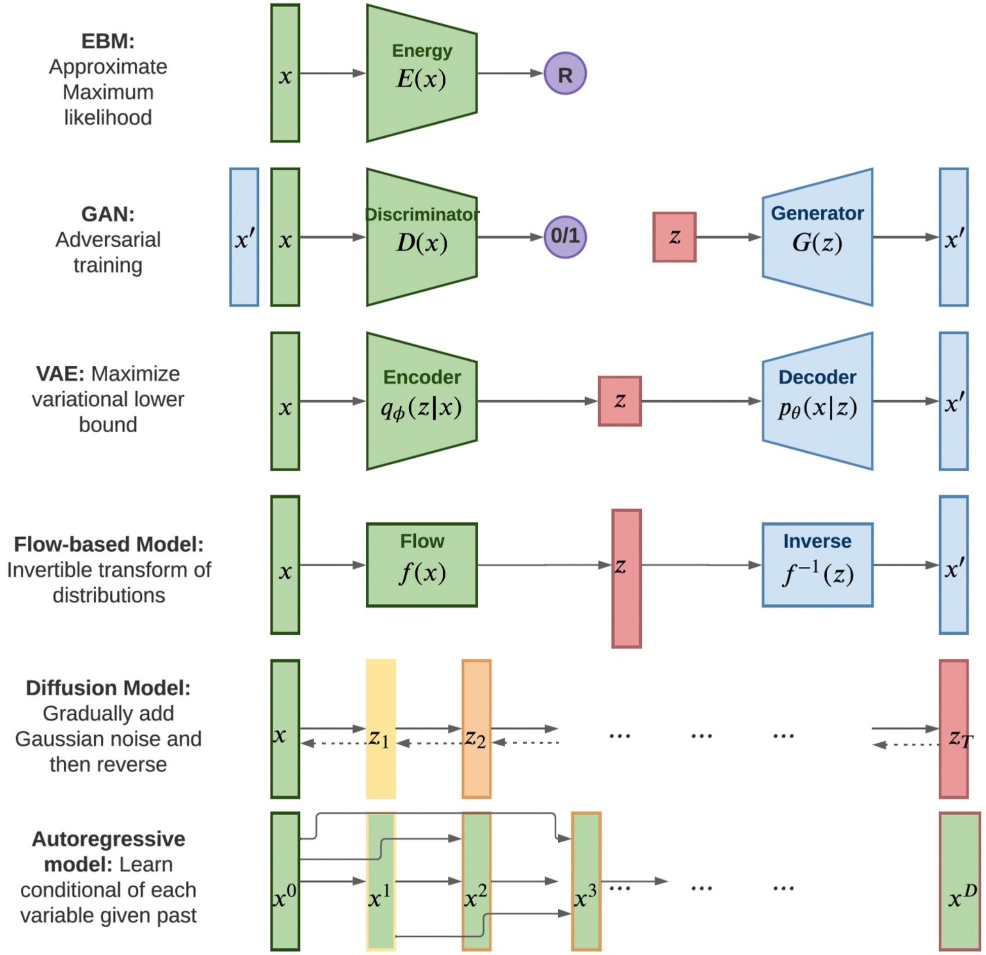
"Cat"



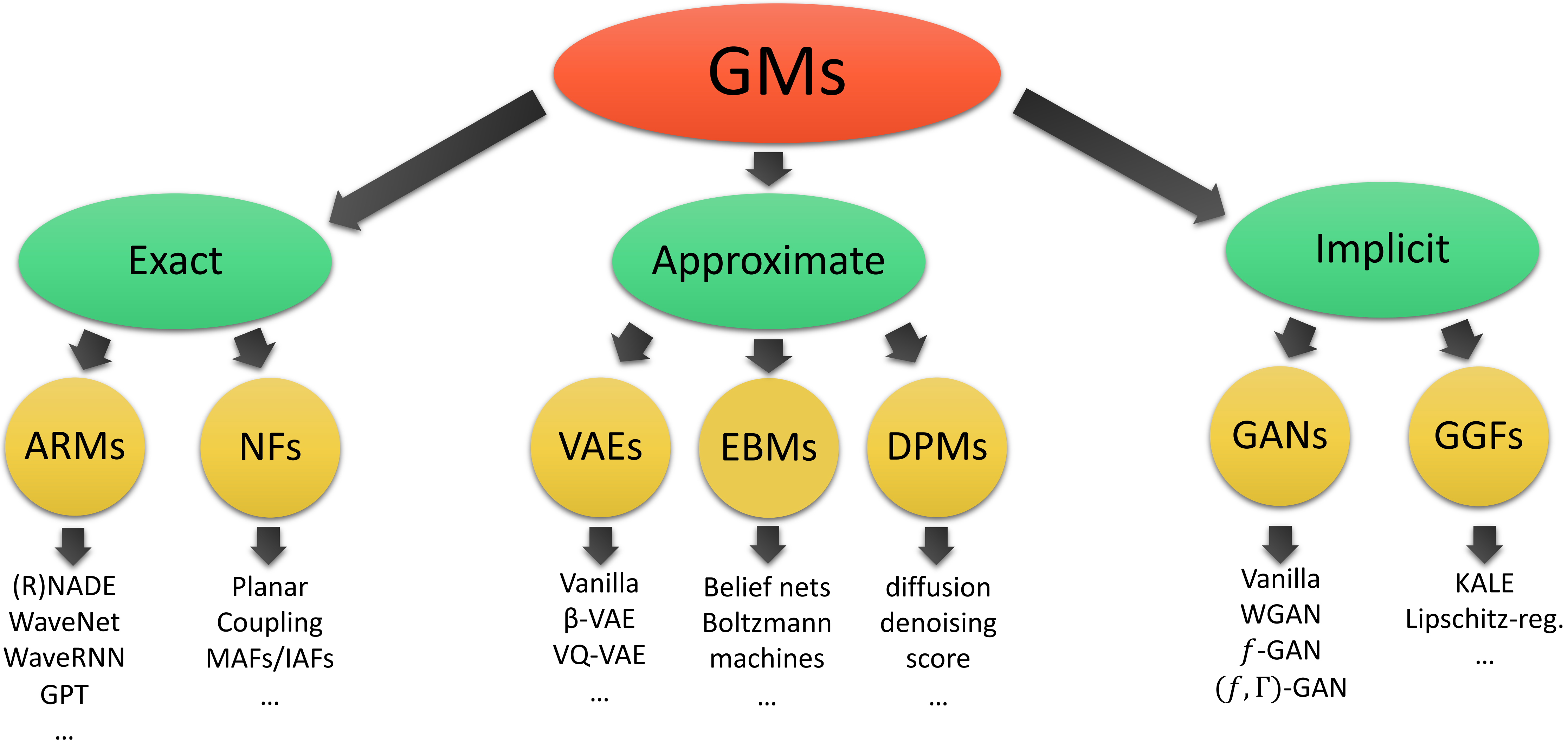
- ✓ **Discriminative Model**
  - ✓ Learn the probability distribution $p(y|x)$
- ✓ **Generative Model**
  - ✓ Learn the probability distribution $p(x)$
- ✓ **Conditional GM**
  - ✓ Learn $p(x|y)$

Families of Generative Models

- ✓Energy-based Models (EBMs)
- ✓Generative Adversarial Nets (GANs)
- ✓Variational Auto-Encoders (VAEs)
- ✓Normalizing Flows (NFs)
- ✓Diffusion Probabilistic Models (DPMs)
- ✓Deep Autoregressive Models (ARMs)



Families of Generative Models

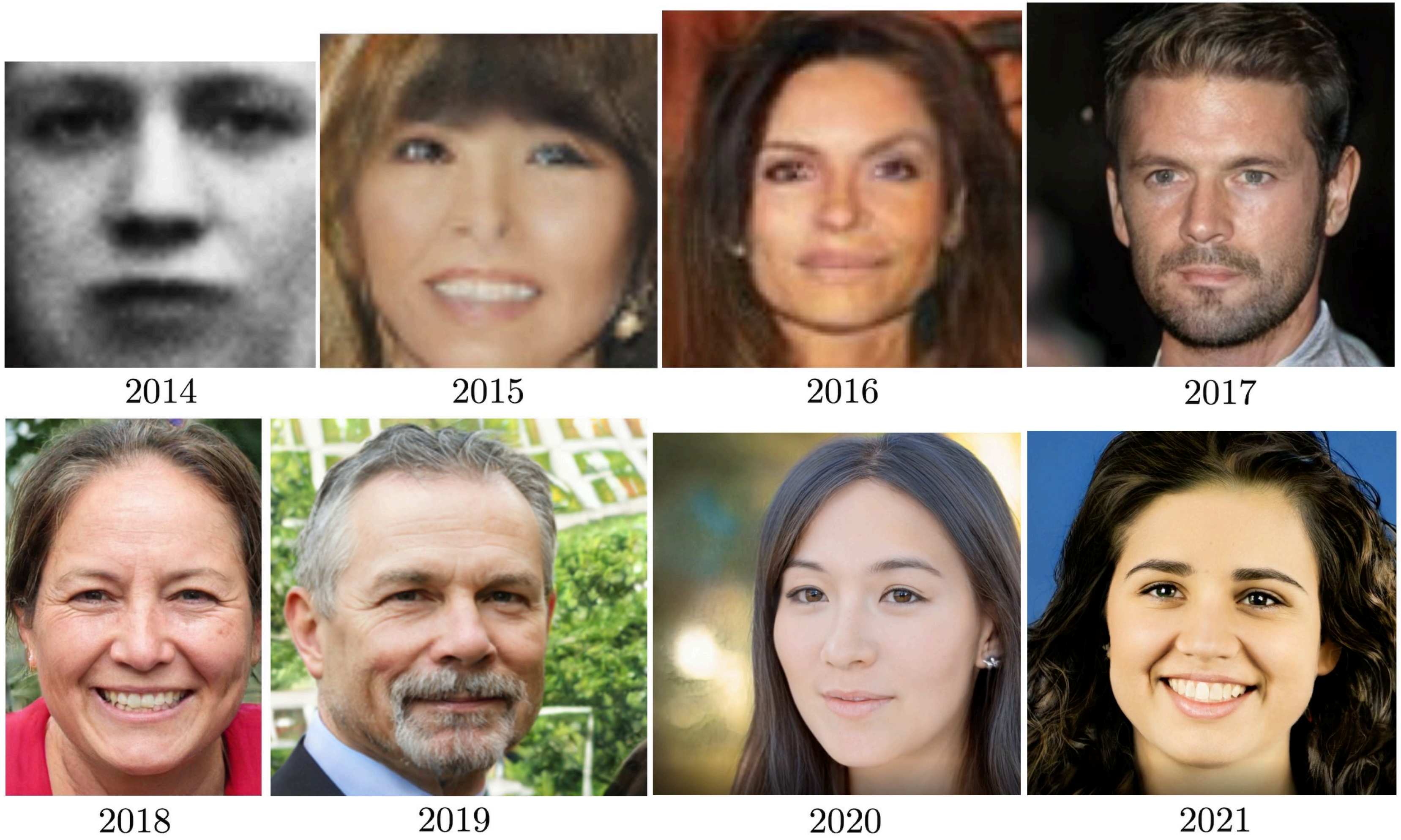


Less known Families of GMs

- ✓Generative Stochastic Networks (GSNs)
- ✓Generative Gradient Flows (GGFs)
- ✓Specific EBMs
  - ✓Deep Belief Networks
  - ✓Deep Boltzmann Machines
  - …
- ✓Generative Flow Networks (GFlowNets)

…

Progress in Image Generation



✓ Face generation: Rapid progress in image quality

Image Super Resolution

- ✓ Several inverse problems can be solved with ***conditional GMs***.
- ✓ *Inverse problems:* From measurements, calculate/infer the causes.

✓ $P(high\ resolution|low\ resolution)$

- ✓ *Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network* - Ledig et al. - CVPR 2017
  - ✓ https://openaccess.thecvf.com/content_cvpr_2017/html/Ledig_Photo-Realistic_Single_Image_CVPR_2017_paper.html

Image Inpainting

- ✓ $\boldsymbol{P(full\ image|masked\ image)}$
- ✓ *DeepFill (v2):* *Free-Form Image Inpainting With Gated Convolution*
  – Yu et al. - ICCV 2019
  - ✓ https://openaccess.thecvf.com/content_ICCV_2019/html/Yu_Free-Form_Image_Inpainting_With_Gated_Convolution_ICCV_2019_paper.html

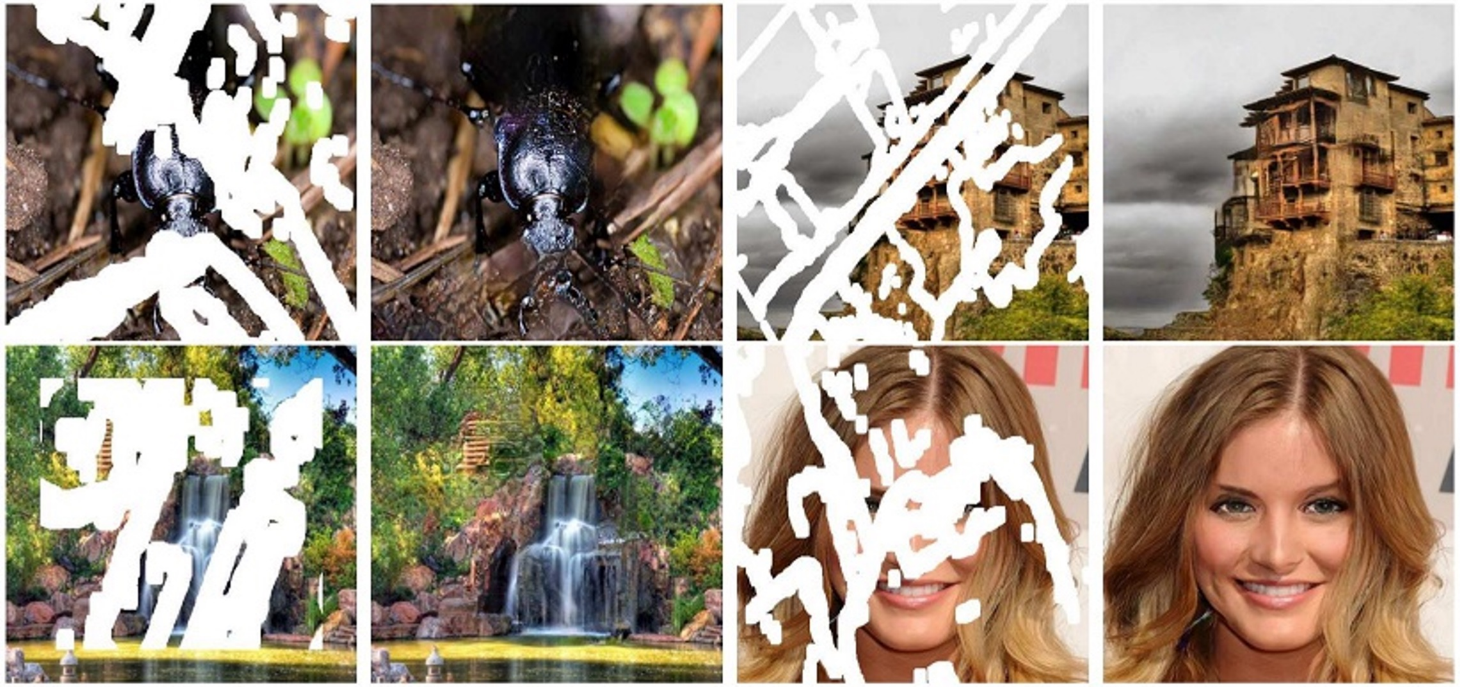


Image Colorization

✓ $P(colored\ image|grayscale\ image)$

✓ *PalGAN: Image Colorization with Palette Generative Adversarial Networks* – Wang et al. - ECCV 2022

  ✓ https://link.springer.com/chapter/10.1007/978-3-031-19784-0_16

Text2Image Translation

Lecture #1

- ✓ ***Recent advancements:***
  - ✓ DALL-E 2
  - ✓ Stable Diffusion
  - ✓ Imagen
  - ✓ GLIDE
  - ✓ Midjourney

✓ $\boldsymbol{P(image|text)}$

Théâtre D'opéra Spatial by Jason Allen and Midjourney

OpenAI's DALL-E 2

Lecture #1

- ✓Text → Text embedding → Image embedding → low resolution image → medium resolution image → high resolution image

- ✓$P(high\,res\,image|text\,caption) = P(image\,emb|text\,caption)$ x $P(high\,res\,image|image\,emb)$

- ✓*Hierarchical Text-Conditional Image Generation with CLIP Latents* - Ramesh et al. - 2022
  - ✓https://cdn.openai.com/papers/dall-e-2.pdf

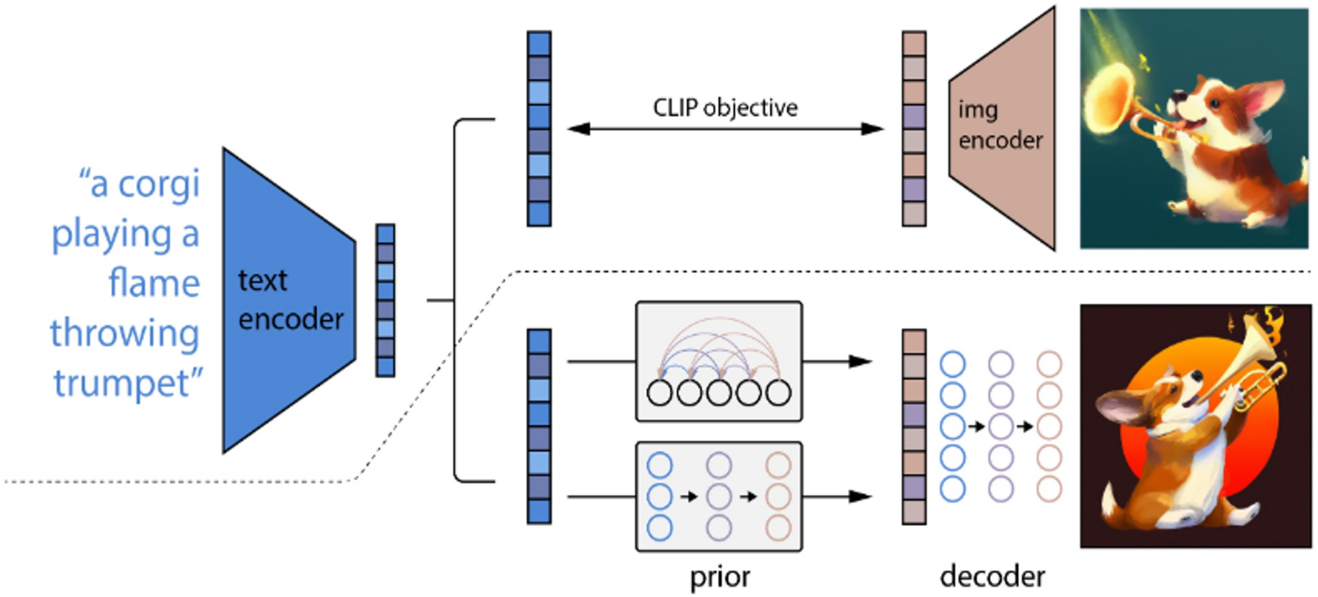
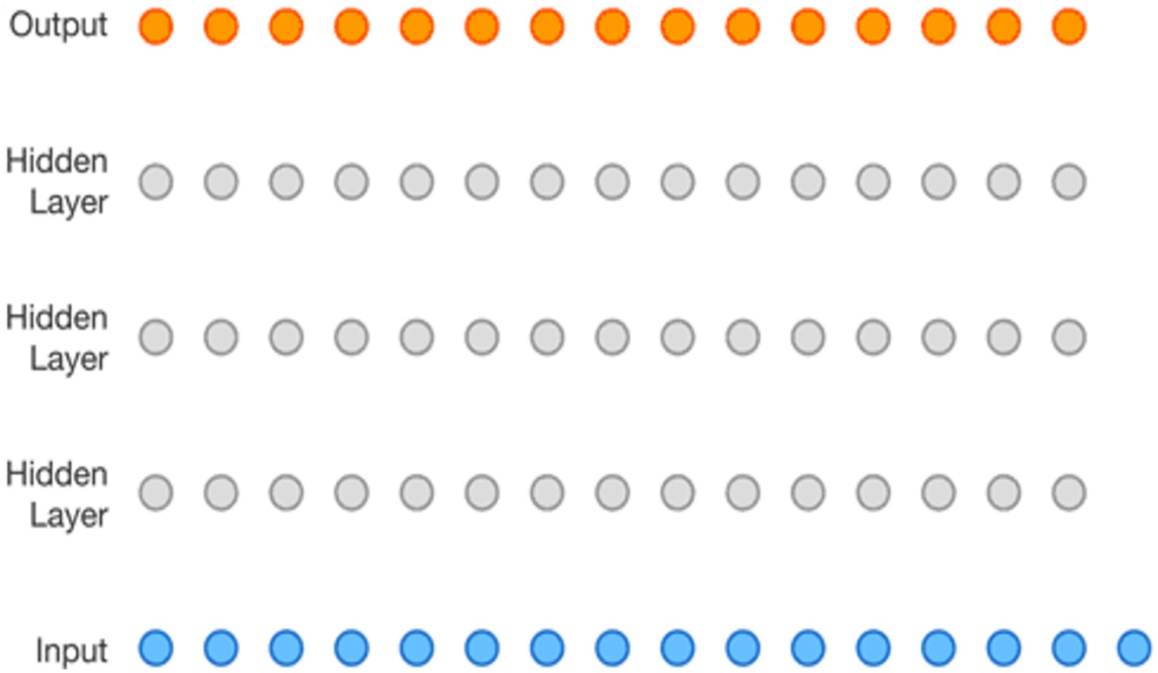


Image2Image Translation

- ✓ *Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks* (CycleGAN) – Zhu et al. - ICCV 2017
  - ✓ https://openaccess.thecvf.com/content_iccv_2017/html/Zhu_Unpaired_Image-To-Image_Translation_ICCV_2017_paper.html

Speech & Audio Synthesis

- ✓ $\boldsymbol{P}(x_{t+1}|x_t, x_{t-1}, \dots, text)$
- ✓ WaveNet, WaveRNN, Parallel Wavenet, MelGAN, WaveDiff, …



Text to Speech Synthesis

- Parametric
- Concatenative
- WaveNet

- Unconditional
- Music

van den Oord et al., 2016

(Natural) Language Generation

Lecture #1

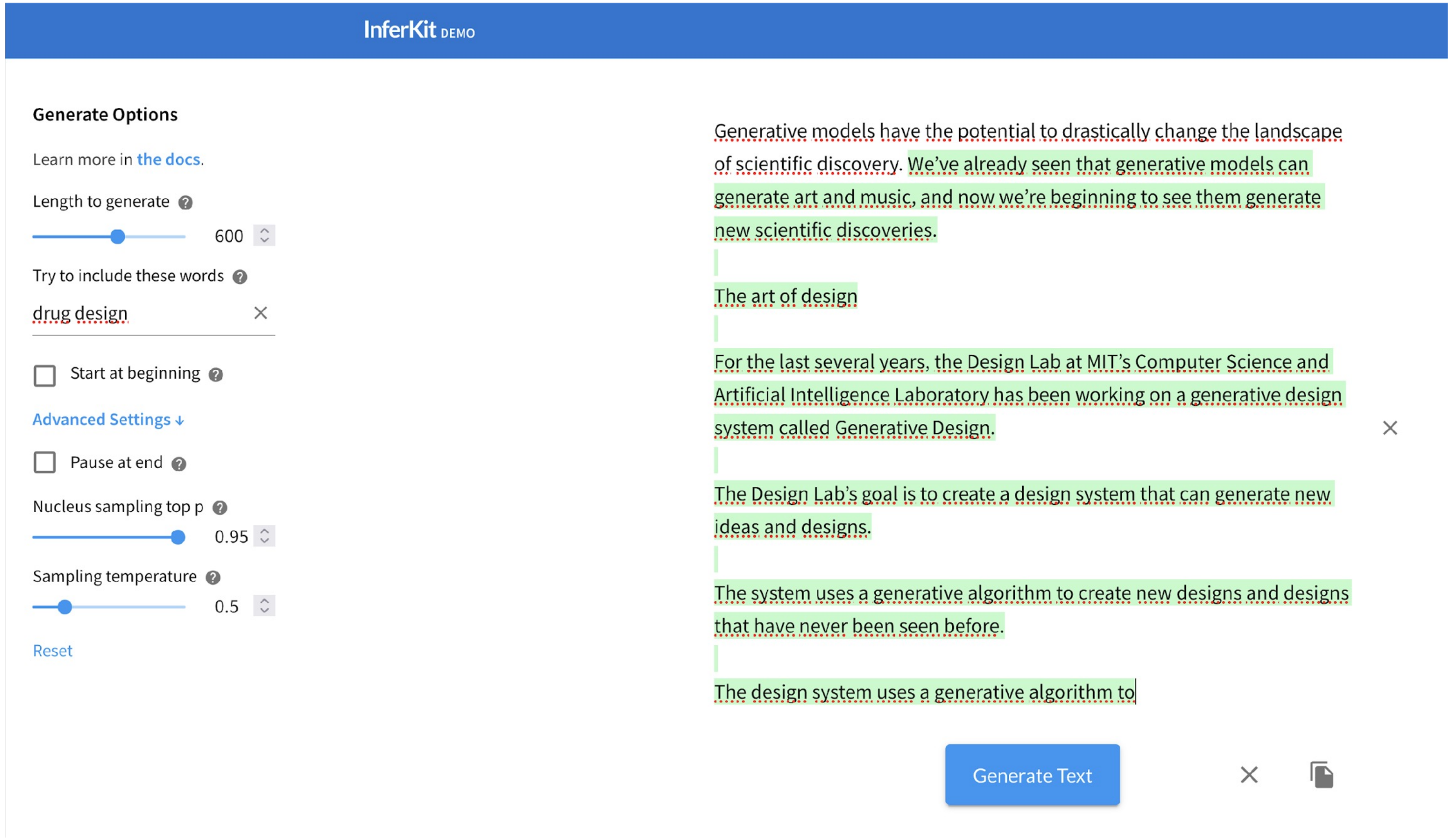
✓ $P(next\ word|previous\ word)$

✓ https://app.inferkit.com/demo

✓ GPT-3

✓ Generative Pre-trained Transformer

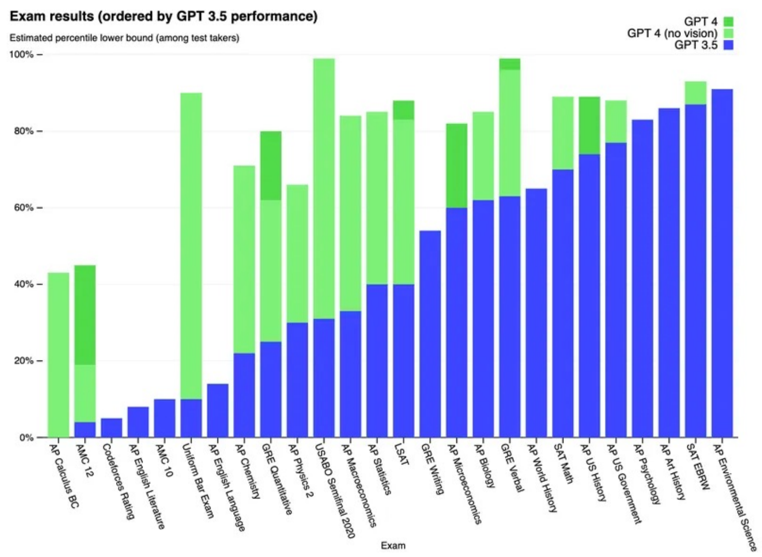
✓ https://deepai.org/machine-learning-model/text-generator



(Natural) Language Generation

- ✓Enormous model size (Trillion parameters?)
- ✓Enormous & diverse training data
- ✓Multimodal capabilities
- ✓Context learning (a.k.a. prompting)
- ✓Reinforcement learning
- ✓Enormous performance
  - ✓Coherence, relevance, proficiency
  - ✓Safety & Ethics
  - ✓Few steps from AGI

OpenAI ChatGPT 4.0

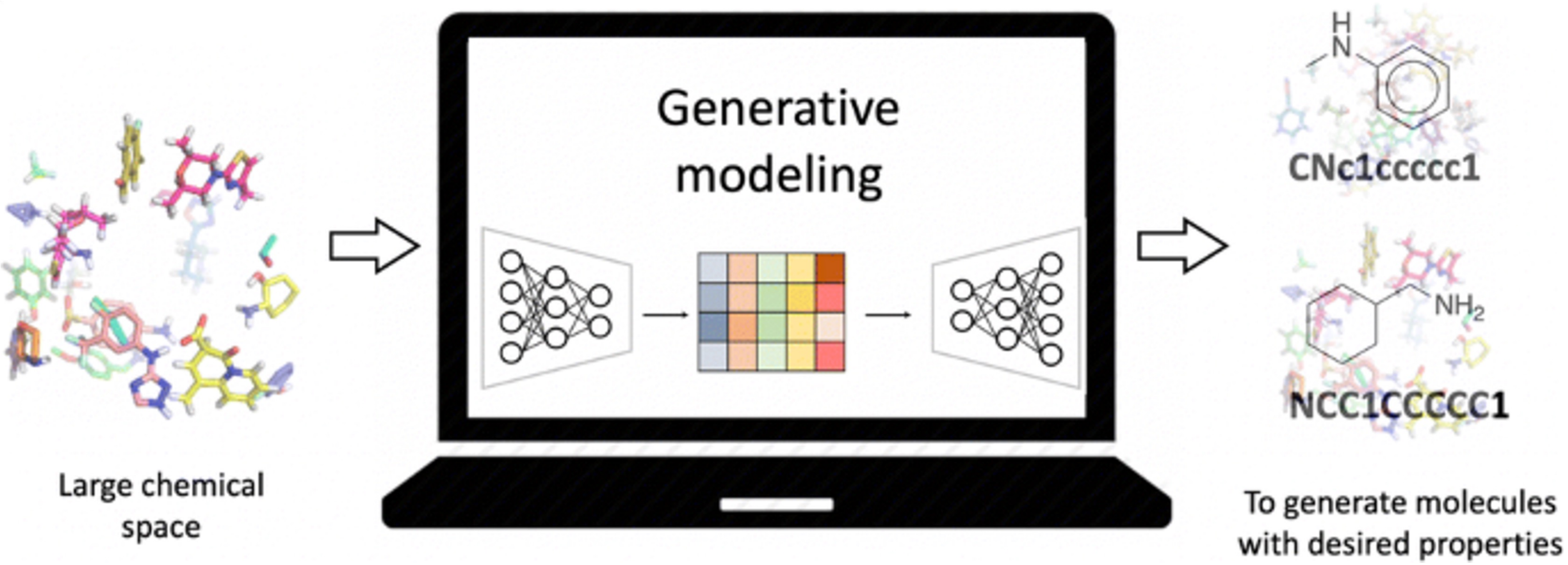


Geometric Design

✓Just meshing around with GPT4 (ESA proposal)

Molecule/Drug/Protein Design

✓ *MolGAN: An implicit generative model for small molecular graphs* – De Cao & Kipf – ICML 2018



Driving forces in GM progress

- ✓<u>Representation learning</u>
  - ✓Leveraging the exponential growth of data & of model's parameters via self-supervised learning
  - ✓Gave raise to the Foundation Models
- ✓<u>Computational resources</u> are also exponentially increasing.
- ✓<u>Better understanding</u> of the models, algorithms act as key enablers.
- ✓<u>Unlocks</u> *human* productivity & *creativity*.
- ✓Ideally, it will <u>accelerate</u> the *scientific discovery process*.

Challenges in GMs

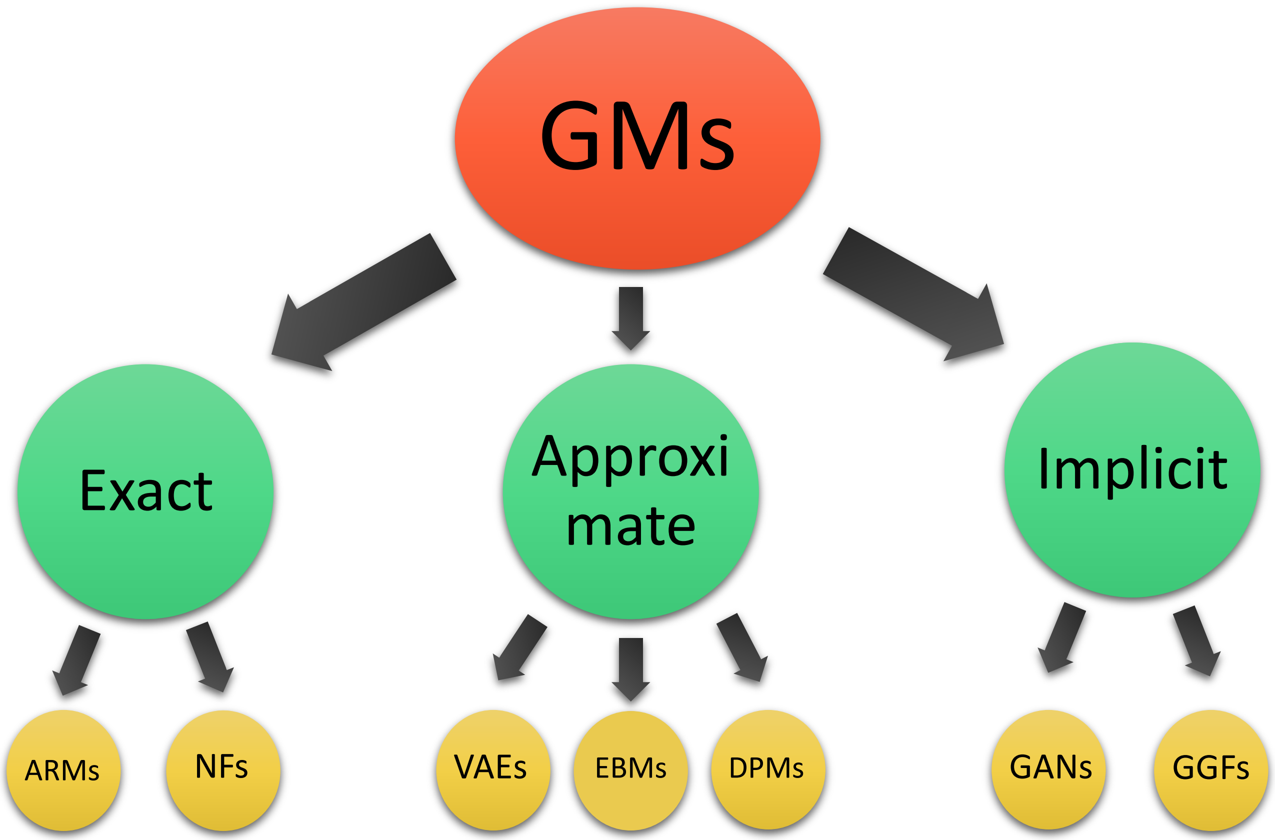
- ✓ ***Representation:*** How do we model the joint distribution of many random variables?
  - ✓ Need compact & meaningful representations
- ✓ ***Learning (a.k.a. quality assessment):*** What is the proper comparison metrics between probability distributions?
- ✓ ***Reliability:*** Can we trust the generated outcomes? Are they consistent?
- ✓ ***Alignment:*** Do they perform according to the input of the user?

Prerequisites

- ✓ Very good knowledge of *probability theory, multivariate calculus & linear algebra*.
- ✓ Intermediate knowledge regarding *machine learning & neural networks*.
- ✓ Proficiency in some *programming language*, preferable *Python*, is required.

Course Syllabus

- ✓Basics in probability theory (1W)
  - ✓Shallow generative models - GMMs (1W)
- ✓Exact (i.e., fully-observed) likelihood
  - ✓AR models (1.5W)
  - ✓Normalizing flows (1.5W)
- ✓Approximate likelihood
  - ✓VAEs (2W)
  - ✓Diffusion/Score-based models (2W)
  - ✓EBMs (1W)
- ✓Implicit
  - ✓GANs (2W)



Logistics

- ✓ *Teaching Assistant:* Michail Raptakis (PhD candidate)
- ✓ *Weekly Tutorial (Friday 10:00-12:00):* Python/PyTorch basics, neural network architectures and training, solve problems to assist with homework, solve selected homework's problems.
- ✓ *Textbook:* ***Probabilistic Machine Learning: Advanced Topics*** by Kevin P. Murphy
  - ✓ https://probml.github.io/pml-book/book2.html
- ✓ Seminal papers will be distributed.

Grading policy

- ✓ *Final Exam* (**30%** of total grade)
  - ✓ Open notes
  - ✓ NO internet
- ✓ 5-6 series of *Homework* (**40%** of total grade)
  - ✓ Mix of theoretical and programming problems
  - ✓ Equally weighted
- ✓ *Project*: paper implementation & presentation (**30%** of total grade)
  - ✓ Implementation: **10%**
  - ✓ Final report: **10%**
  - ✓ Presentation: **10%**

Project

- ✓Select from a given list of papers or propose a paper (which has to be approved)
- ✓*<u>Categories of papers</u>*:
  - ✓Application of deep generative models on a novel task/dataset
  - ✓Algorithmic improvements into the learning, inference and/or evaluation of deep generative models
  - ✓Theoretical analysis of any aspect of existing deep generative models
- ✓Groups of **up to 2 students** per project
- ✓Computational resources might be provided (colab, local GPUs, etc.)

Introduction to Deep Generative Modeling

Lecture #1

HY-673 – Computer Science Dept, University of Crete
Professors: Yannis Pantazis & Yannis Stylianou
TAs: Michail Raptakis & Michail Spanakis