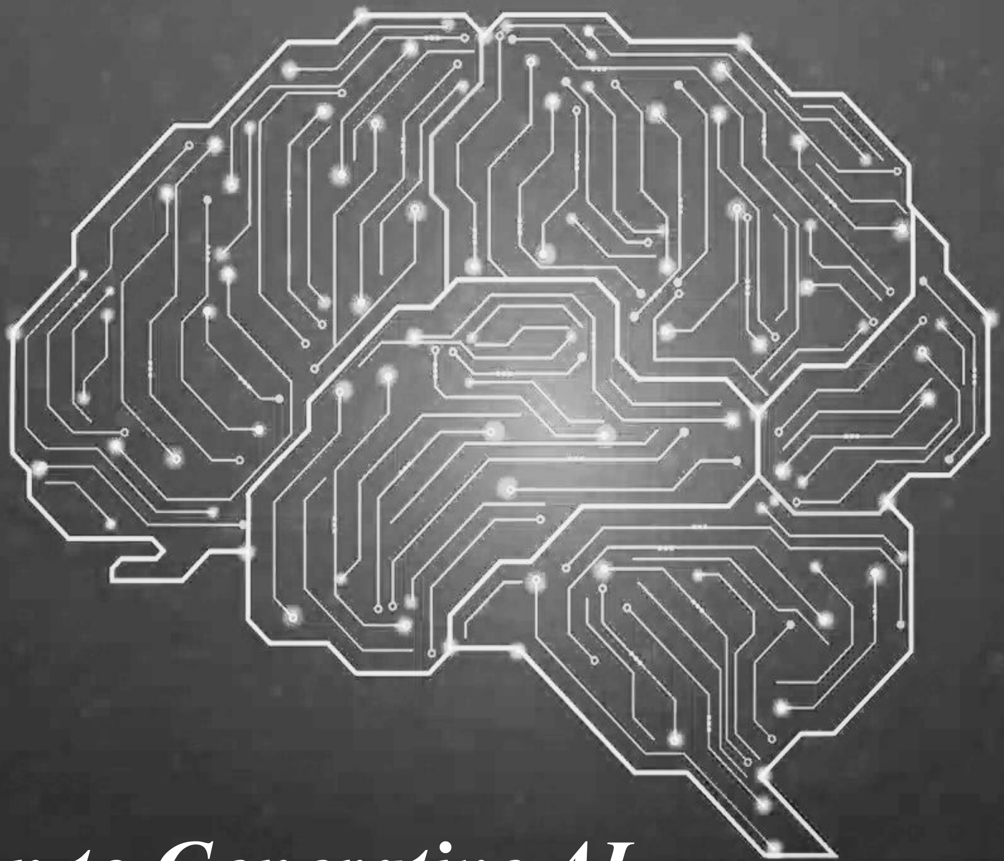


Introduction to Generative AI

Dr. Krupali Donda

Your faculty for this course..

Ph.D. from IJL, CNRS

The French National Centre for Scientific Research (with USA Airforce)

Research Fellow, IIT Gandhinagar

Research and teaching areas: Scientific Machine Learning, Metasurface, Electronics and Communication

Ardent traveler and explorer (14 countries) ☺

Theory Marks (100)

End Semester Examination : 100 Marks

Continuous Evaluation : 50 Marks (Mid sem.=30marks, Case studies=10marks, Class notes=10marks)

Practical Marks (50)

Practical Assessment : 25 Marks

Viva : 25 Marks

Syllabus

Chapter 1: Introduction to Generative AI

- ➔ Overview of Generative AI
- ➔ Generative vs Discriminative AI
- ➔ Categories of Generative Models (Explicit vs Implicit)
- ➔ Applications
- ➔ Ecosystem: OpenAI, Hugging Face, Stability.ai
- ➔ Responsible and Ethical Use

Chapter 2: Variational Autoencoders and Diffusion Models

- ➔ Autoencoder Architecture Recap
- ➔ Variational Autoencoders (VAE)
- ➔ Diffusion Models: DDPM, DDIM
- ➔ Stable Diffusion and Latent Diffusion
- ➔ Comparison: VAEs vs GANs vs Diffusion

Chapter 3: Generative Adversarial Networks (GANs)

- GAN Architecture: Generator & Discriminator
- Loss Functions and Training Instability
- GAN Variants: DCGAN, cGAN, CycleGAN, StyleGAN
- Evaluation Metrics: IS, FID
- Applications

Chapter 4: Transformers and Attention Mechanisms

- Self-Attention and Multi-Head Attention
- Transformer Encoder-Decoder Architecture
- Positional Encoding
- Comparison with CNNs and RNNs
- Foundation for LLMs

Chapter 5: Large Language Models (LLMs) and Foundation Models

- ➔ Pretrained Language Models: GPT, BERT, T5, LLaMA
- ➔ In-Context Learning and Fine-Tuning
- ➔ Instruction Tuning and RLHF
- ➔ LLM Ecosystem: OpenAI, Anthropic, Meta
- ➔ Alignment, Bias, and Hallucination

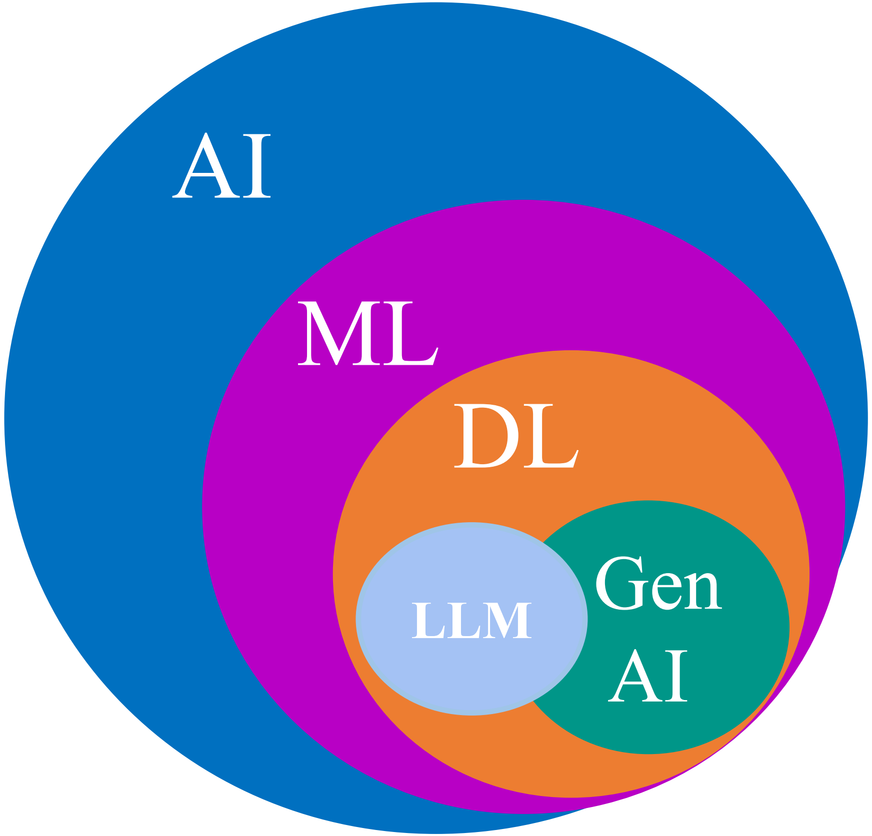
Chapter 6: Prompt Engineering and Retrieval-Augmented Generation (RAG)

- ➔ Prompt Design Techniques: Zero-shot, Few-shot, CoT
- ➔ Prompt Failures and Refinement
- ➔ Retrieval-Augmented Generation (RAG)
- ➔ Embeddings and Vector Search (FAISS, Pinecone)
- ➔ LangChain and LlamaIndex Overview
- ➔ Prompt Injection and Security

"I imagine a world in which AI is going to make us work more productively, live longer, and have cleaner energy."

Fei-Fei Li
Professor of Computer Science
Stanford University

Generative AI



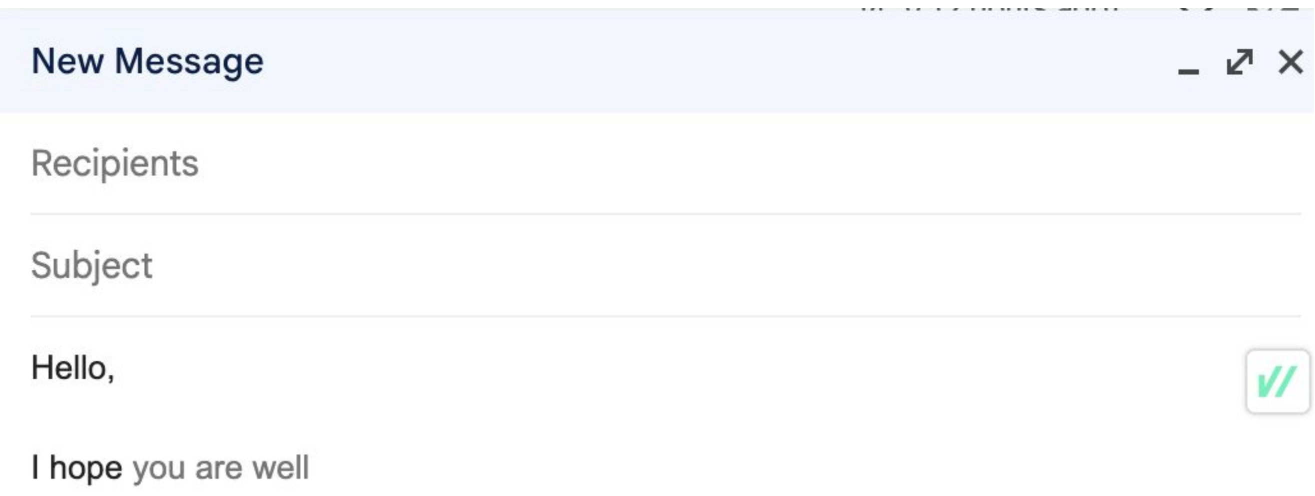
What is Generative AI?

Generative	Create new content (text, image, video, audio)
+	
Artificial Intelligence	Automatically using a Computer Programme

Generative AI is a subfield of artificial intelligence focused on the creation of models that can synthesize novel data instances that are statistically similar to a given training distribution.

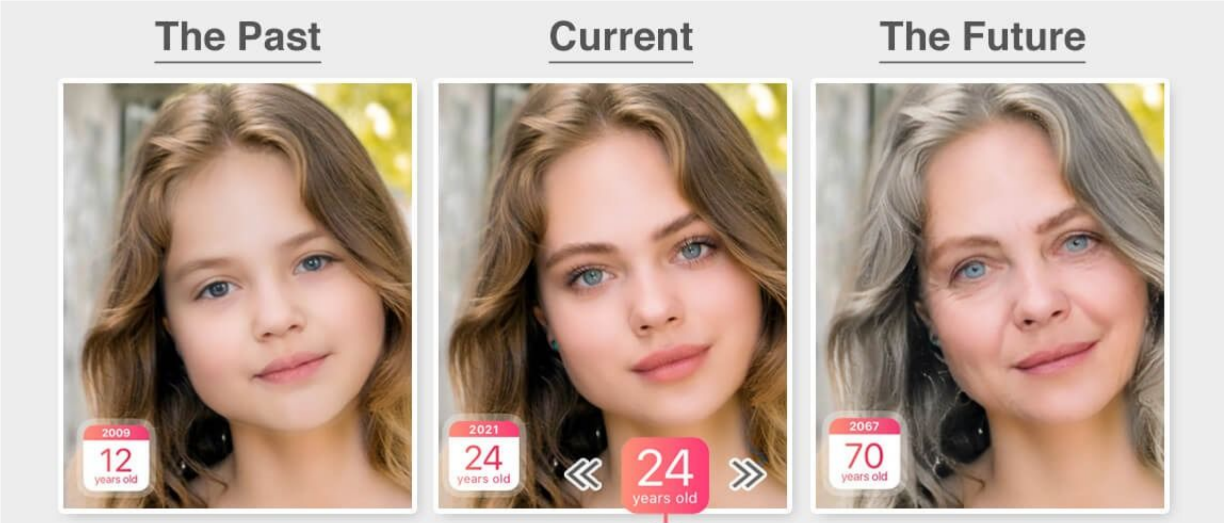
Generative AI is not a new topic!

Predictive Text & Autocomplete (Gmail Smart Compose, 2018)



Generative AI is not a new topic!

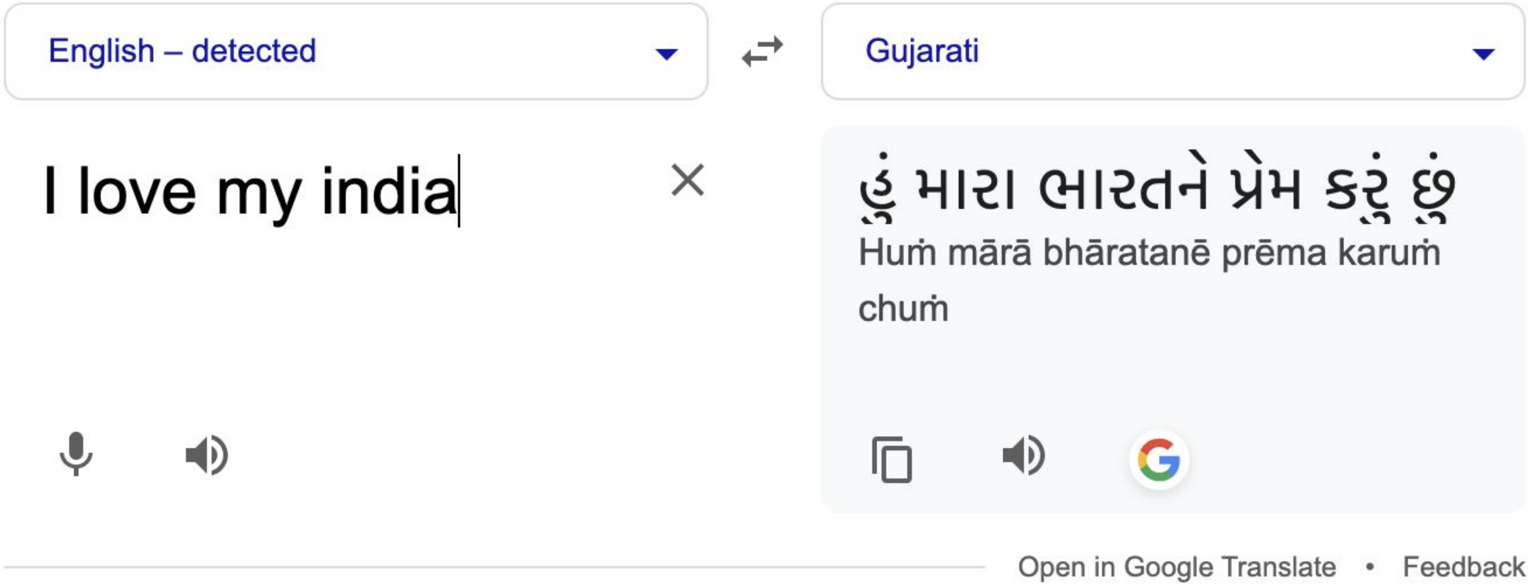
Photo Filters and Face Apps (Snapchat, FaceApp, 2016)



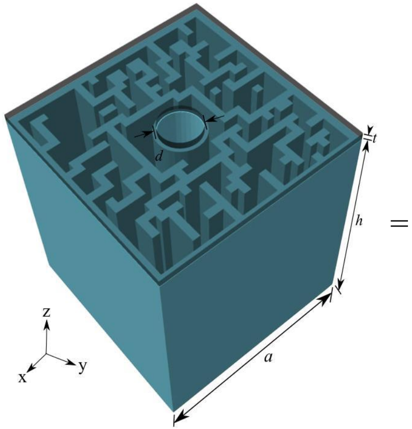
How It's Generative: Uses generative models to alter or create new versions of your face (e.g., aging, smiling, gender swap).

Generative AI is not a new topic!

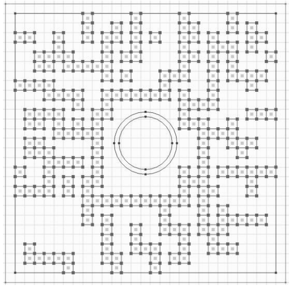
Statistical Machine Translation (pre-2015 Google Translate)



Gen AI for Metamaterial Design (2021)



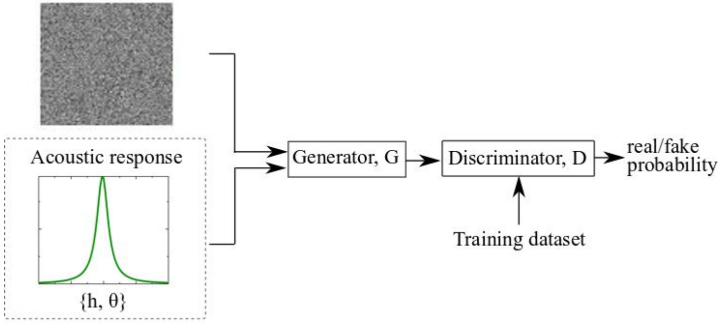
2D Pattern



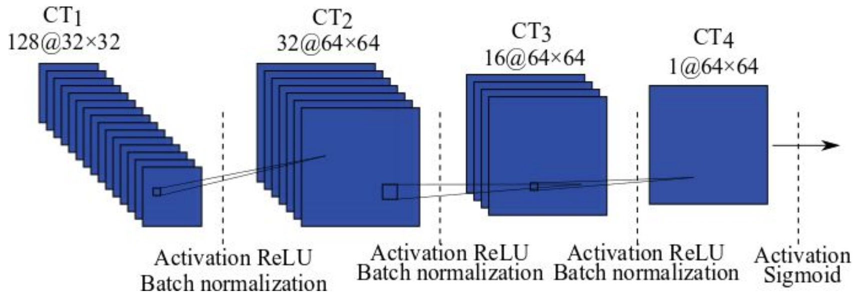
+

1D properties

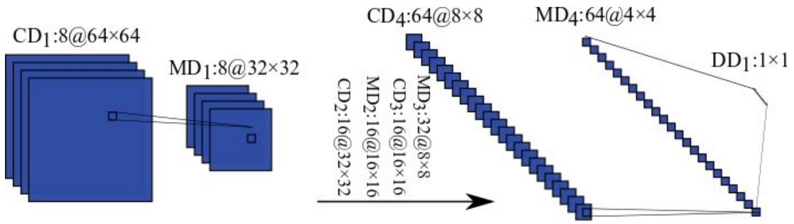
{h, θ}



Generator



Discriminator



Than What happened?

ChatGPT passes exams from law and business schools

By Samantha Murphy Kelly, CNN Business
4 minute read · Updated 1:35 PM EST, Thu January 26, 2023

TECH

OpenAI announces GPT-4, claims it can beat 90% of humans on the SAT

PUBLISHED TUE, MAR 14 2023•1:42 PM EDT | UPDATED TUE, MAR 14 2023•2:32 PM EDT

Kif Leswing
@KIFLESWING

SHARE

You can also ask ChatGPT..

Write my introduction for Linkedin. My name Krupali. Did my phd from CNRS, France in acoustic metamaterials. Love travelling and exploring. I love writing on Linkedin. I am Assistant Professor. Teaching Generative AI…

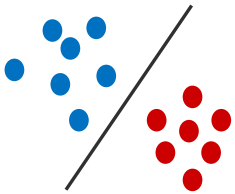
Write the simplest Python code to implement a neural network on the Auto MPG dataset and predict the MPG.

It’s a monsoon time here in Gujarat. Write a poem in Gujarati for the season. Also, don’t forget to mention tea in the poem.


Time it took to reach 100 million monthly users:
App
Months To Reach 100M Global MAU
ChatGPT 2
TikTok 9
Instagram 30
Pinterest 41
Uber 70
Google Translate 78
Source: UBS
@TheRundownAI

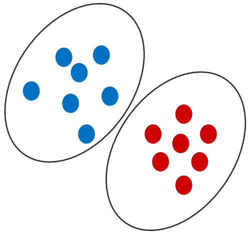
Types of DL: Discriminative and Generative Model

Discriminative Model



- Learn Conditional Probability
- Focus: Boundary b/w classes
- Use: Classification and Prediction
- Mostly trained on labelled dataset.
- Learn the relationship between data and corresponding labels.
- Example: Logistic regression, SVM

Generative Model



- Learn Joint Probability
- Focus: Distribution of each class
- Use: Classification+Generation
- Generate new data based on the probability distribution of original data.
- Example: GAN, VAE

Example

Discriminative Method

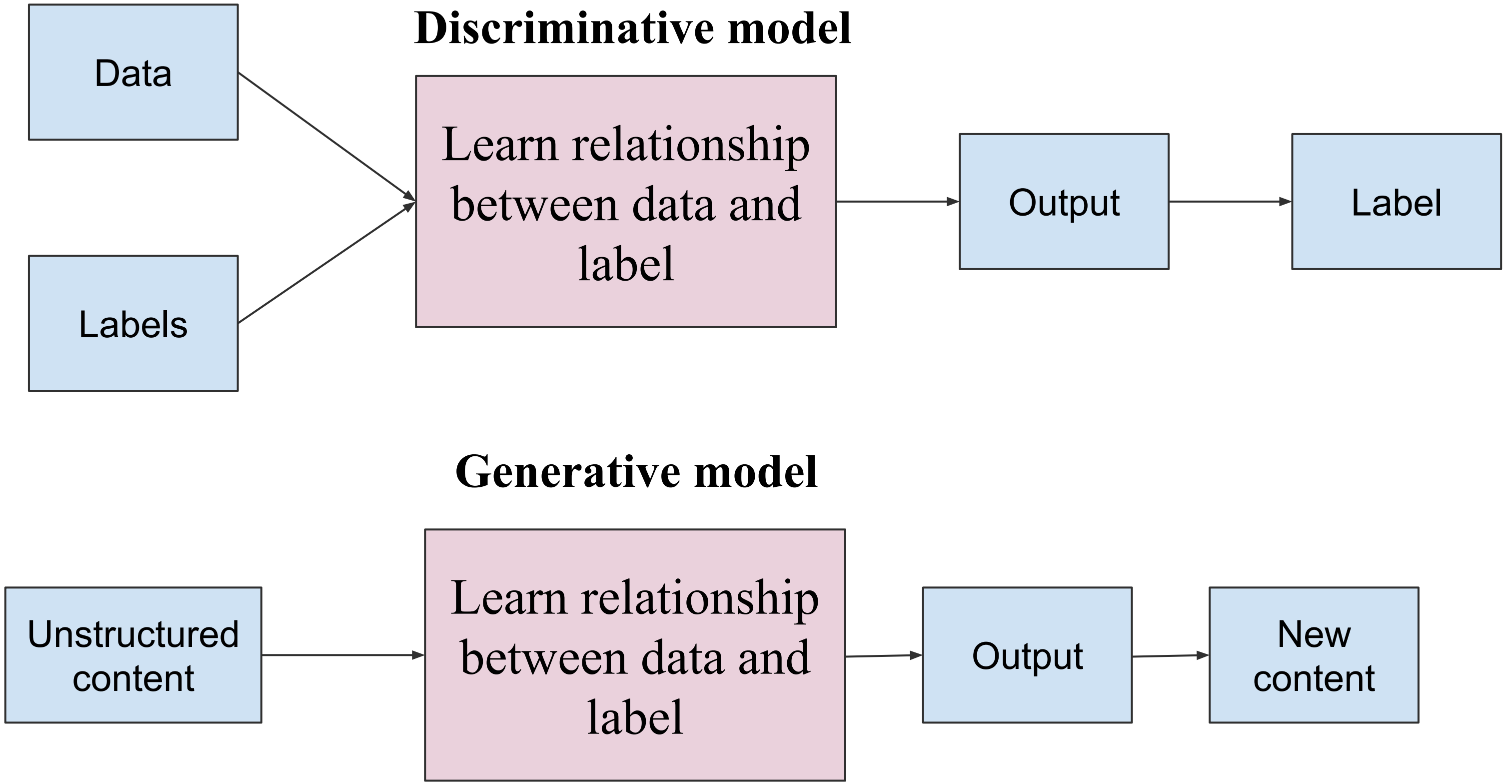
Classify

Discriminative Model (classify as white or red rose)

Generative Method

Generate

Generative Model (generate red rose)


Data
Discriminative model
Learn relationship between data and label
Output
Label
Labels
Generative model
Unstructured content
Learn relationship between data and label
Output
New content

Types of Generative AI (Based on Input Modality:Image)

Input: Image

Output: Text

- Image caption
- Image description

Output: Image

- High resolution
- Image enhancement
- Style transfer

Output: Video

- Image-to-video generation
- Motion prediction
- Cinemagraph creation
- Time-lapse generation

Types of Generative AI (Based on Input Modality: Text)

Input: Text

O/P: Text

- Translation
- Summarize
- Question-answering
- Creative writing

O/P: Image

- Postcard creation
- Image generation

O/P: Audio

- Text-to-speech
- Music synthesis
- Voice cloning

O/P: Decision

- Game strategy
- Risk Assessment

Types of Generative AI (How the Model Learns the Data Distribution)

1. Explicit Density Models
2. Implicit Density Models

Explicit Density Models

- ➢ Generative models that explicitly represent and compute the probability distribution p(x) for the data.

  (They don’t just generate new data samples, they also know exactly how likely or probable each sample is under the learned distribution.)

- ➢ Example: Food Delivery App

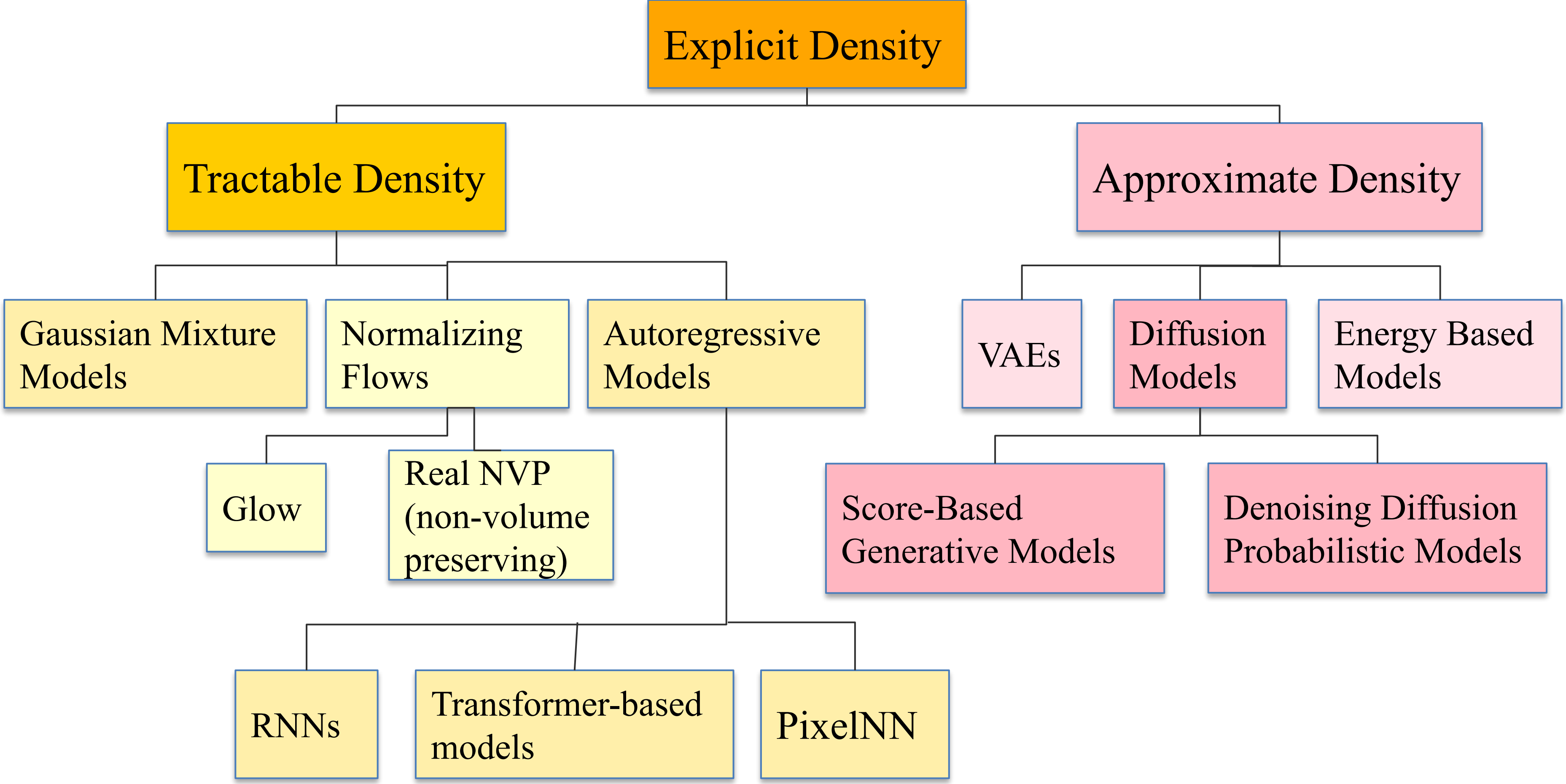
- The goal is to model or approximate the probability distribution *Pmodel*(*x*) of data *x*.
- Trained by maximizing the likelihood of the data,

$$\theta^* = \arg\max_{\theta} \sum_{i=1}^{N} \log p_{\text{model}}(x_i; \theta)$$

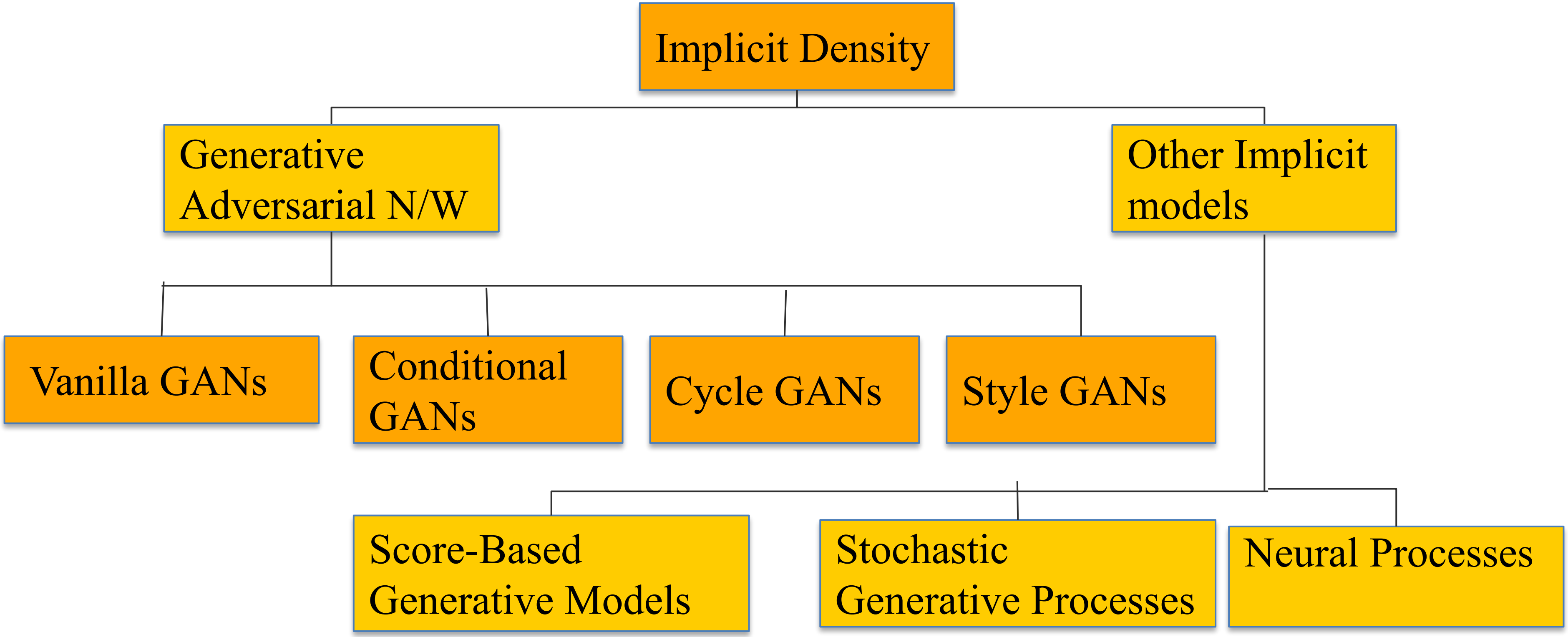
xi = Each data point in the training set
pmodel(xi;θ) = Probability the model assigns to x_i, given parameters θ

log: Log-likelihood (used for stability and simplification)

θ*: Optimal parameters that maximize the total log-likelihood over all N data points


Explicit Density
Tractable Density
Approximate Density
Gaussian Mixture Models
Normalizing Flows
Autoregressive Models
VAEs
Diffusion Models
Energy Based Models
Glow
Real NVP (non-volume preserving)
Score-Based Generative Models
Denoising Diffusion Probabilistic Models
RNNs
Transformer-based models
PixelNN

Implicit Density Models



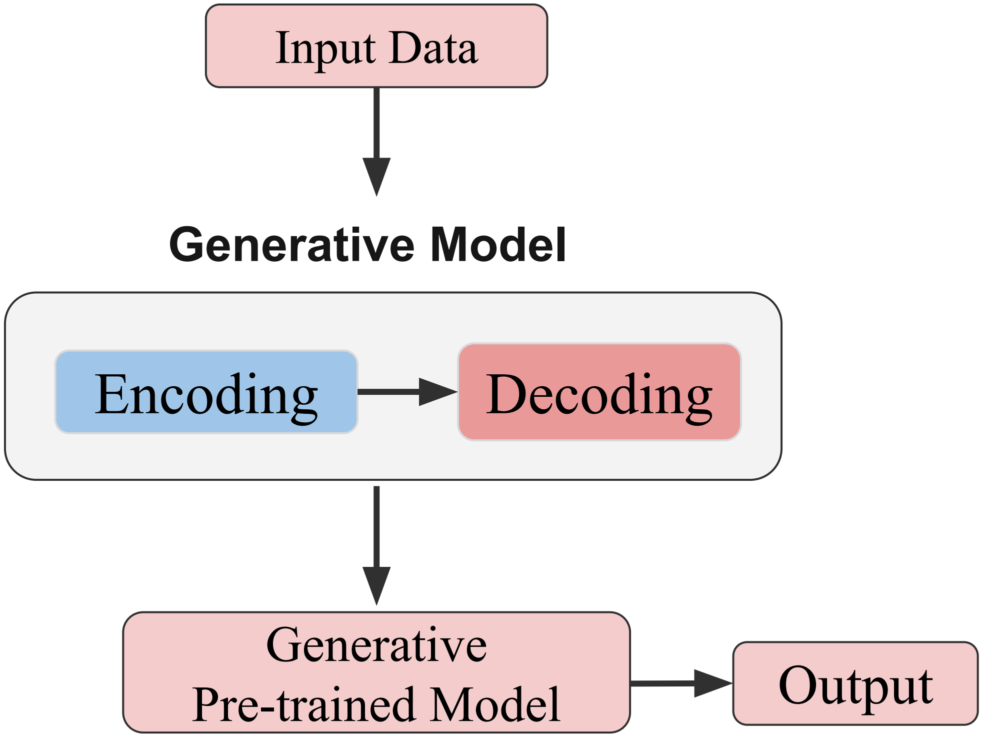
Hallucination

Hallucinations in generative AI occur when the model produces plausible-sounding but false or misleading information. These are not intentional errors but arise due to limitations in data or training.

Hallucination

- ⚠️ The model is not trained on enough data
- ⚠️ The model is trained on noisy or dirty data
- ⚠️ The model is not given enough context
- ⚠️ The model is not given enough constraints
- ⚠️ Overgeneralization from training
- ⚠️ Bias or Gaps in Training Data
- ⚠️ Outdated Information

High-level Overview: How Gen AI works?



Training

- Large dataset
- Billions of parameters
- Unsupervised learning

Encoding: Converts input data into a compressed latent representation (Captures semantic structure and meaning.)

Overgeneralization

Generating outputs that are too broad, incorrect, or stereotyped because it applies patterns too widely from its training data.

The nature or general characteristics of European people can vary greatly because Europe is a continent of **wide cultural, linguistic, and historical diversity**. However, some general observations—while not true for everyone—can be made about certain shared values and behaviors:

🌍 General Traits Often Associated with Europeans

(These vary across regions and individuals, and should not be taken as stereotypes.)

---

1. Value for Individualism and Privacy

- In many European countries, especially in Northern and Western Europe (like Germany, the UK, the Netherlands, and Scandinavia), people tend to value **personal space, independence, and self-expression**.
- They may not appear warm at first but are often **loyal and sincere in relationships**.

Applications: Gen AI (2025)

Text	Code	Image	Speech
- Marketing content - Sales emails - Chat - Customer support - Writing - Note-taking - Translation	- Code generation - Code explanation - Text to SQL/API - App/UI builder	- Image generation from text - Product mockups - Ad creatives - Meme creation - Medical & industrial imaging	- Voice synthesis - Conversational avatars - Audiobook narration - Multilingual voice translation

Applications: Gen AI

Video	3D	Science
- Video generation from text/image/audio	- 3D scene generation	- Protein structure prediction
- Animation	- Game asset design	- Molecular generation
- Ads & explainer videos	- Architecture & product design	- Materials discovery
- Auto-video editing	- AR/VR asset creation	- Biological simulation

Applications: Gen AI

Automation

- AI agents (task automation)
- Report generation
- RPA (Robotic Process Automation)
- Custom workflows

Gen AI Ecosystem

Feature/Aspect	OpenAI	Hugging Face	Stability AI
Focus	Language + Multimodal AI	Model sharing + NLP tools	Stable Diffusion
API access (commercial + fremium)	API acces (commercial +)	BLOOM, BERT DistilBERT	Stable DF
Community contribution	Limited	Very strong	Growing
Deployment Tools	ChatGPT, API	Transformers, Spaces AutoTrain	DreamStudio, APIs

Ethics of AI

"Getting AI governance right is one of the most consequential challenges of our time. It requires mutual learning, grounded in the lessons and best practices emerging from diverse jurisdictions across the globe."

Global AI Ethics and Governance Observatory, UNESCO

Ten Core Principles

1 PROPORTIONALITY AND DO NO HARM

The use of AI systems must not go beyond what is necessary to achieve a legitimate aim. Risk assessment should be used to prevent harms which may result from such uses.

BREAKTHROUGH PROVISION

No use of AI for social scoring or mass surveillance

The Recommendation is the first international normative instrument that contains a provision against using AI systems for social scoring and mass surveillance purposes.

Ten Core Principles

2 SAFETY AND SECURITY

Unwanted harms (safety risks) as well as vulnerabilities to attack (security risks) should be avoided and addressed by AI actors.

KEY CONCEPT

AI actors and the AI life cycle

AI actors are any actors (natural or legal persons) involved in any stage of the AI life cycle, ranging from research, design, and development to deployment and use, including maintenance, operation, trade, financing, monitoring and evaluation, end-of-use, disassembly and termination.

3 RIGHT TO PRIVACY AND DATA PROTECTION

Privacy must be protected and promoted throughout the AI lifecycle. Adequate data protection frameworks should also be established.

CASE 1

Up close and personal

The data that we share online can have an impact on our individual privacy, often unbeknownst to us. Individuals' behaviour online, including abstract information such as patterns of social media likes and scrolling speeds, may be modelled and used as a basis for targeted advertising or behavioural manipulation.

4 MULTI-STAKEHOLDER AND ADAPTIVE GOVERNANCE AND COLLABORATION

International law and national sovereignty must be respected in the use of data, meaning States can regulate the data generated within or passing through their territories. Additionally, participation of diverse stakeholders is necessary for inclusive approaches to AI governance.

Ten Core Principles

5 RESPONSIBILITY AND ACCOUNTABILITY

AI systems should be auditable and traceable. There should be oversight, impact assessment, audit and due diligence mechanisms in place to avoid conflicts with human rights norms and threats to environmental wellbeing.

CASE 2 Automated rejection

When applying for a loan, it is possible that your bank uses AI to make an automated assessment of your finances and determine if your application will be approved. If these decisions are taken without human oversight and accountability, the consequences can be significant. First, an AI system that is not checked by a human may make a mistake. Second, there is no clear appeals process if there is nobody who can take ultimate responsibility for the decision.

Ten Core Principles

6 TRANSPARENCY AND EXPLAINABILITY

The ethical deployment of AI systems depends on their transparency and explainability. For example, people should be made aware when a decision is informed by AI. The level of transparency and explainability should be appropriate to the context, as there may be tensions between transparency and explainability and other principles such as privacy, safety and security.

KEY CONCEPT

Explainability

The term 'black box' has been used to describe AI wwsystems which are opaque and difficult to interpret. 'Explainability' requires that the logic behind algorithmic decision-making can be fully interpreted by experts and that this logic can be explained to users in accessible language.

Ten Core Principles

7 HUMAN OVERSIGHT AND DETERMINATION

Member States should ensure that AI systems do not displace ultimate human responsibility and accountability.

8 SUSTAINABILITY

AI technologies should be assessed against their impacts on 'sustainability', understood as a set of constantly evolving goals including those set out in the UN's Sustainable Development Goals.

Ten Core Principles

9 **AWARENESS AND LITERACY**

Public understanding of AI and data should be promoted through open and accessible education, civic engagement, digital skills and AI ethics training, media and information literacy.

Ten Core Principles

10 FAIRNESS AND NON-DISCRIMINATION

AI actors should promote social justice, fairness, and non-discrimination while taking an inclusive approach to ensure AI's benefits are accessible to all.

Modified from Burlina P., Joshi, N., Paul, W., Pacheco K. D., and Bressler, N.M. (2021). Addressing artificial intelligence bias in retinal diagnostics.

KEY CONCEPT Machine learning

Machine learning is an application of AI that enables systems to learn from data and to improve without being explicitly programmed, thus improving their accuracy over time.