Week 0e: Generative AI

The Creation Challenge

Machine Learning for Smarter Innovation

BSc-Level Course

October 6, 2025

Outline

Traditional ML:
"What is this?"

- Email: spam or not?
- Image: cat or dog?
- Text: positive sentiment?
- Patient: high risk?

Limitation: Analysis only

Generative AI:
"Create something new"

- Generate: realistic images
- Write: coherent articles
- Compose: original music
- Design: novel molecules

Power: Creation & innovation

The fundamental shift: from understanding existing data to creating new possibilities

Discriminative Models

Learn: $P(y|x)$
"Given input x, what's the label y?"

Examples:

- Logistic regression
- Random Forest
- Neural networks (classification)
- SVM

Goal: Decision boundaries

Generative Models

Learn: $P(x)$ or $P(x, y)$
"What does the data distribution look like?"

Examples:

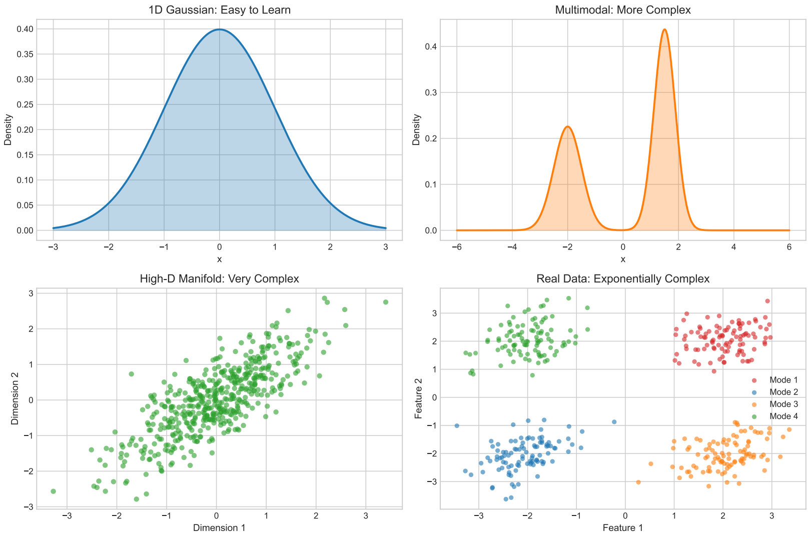
- Gaussian Mixture Models
- Variational Autoencoders
- GANs
- Diffusion models

Goal: Data generation

Discriminative: "Is this a cat?" — Generative: "Draw me a cat"

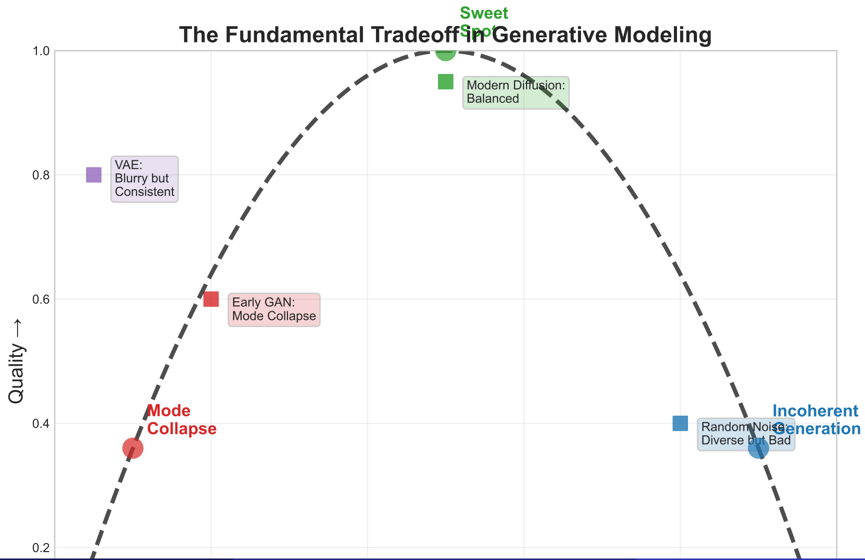
The Hard Problem

Why Generation is Fundamentally Difficult



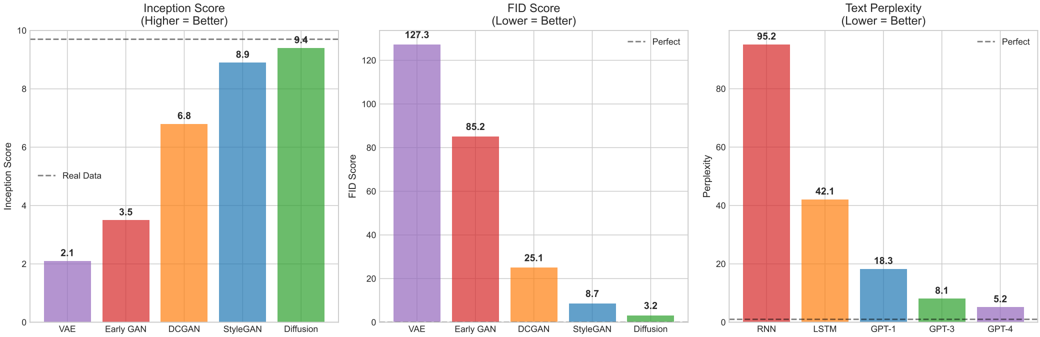
The Fundamental Tradeoff

Quality vs Diversity Dilemma



Measuring Generation Quality

Metrics for Evaluating Generative Models



Inception Score (IS)

- Range: 1-1000+
- Higher = better
- Quality & diversity
- $IS = \exp(E[KL(p(y|x)||p(y))])$

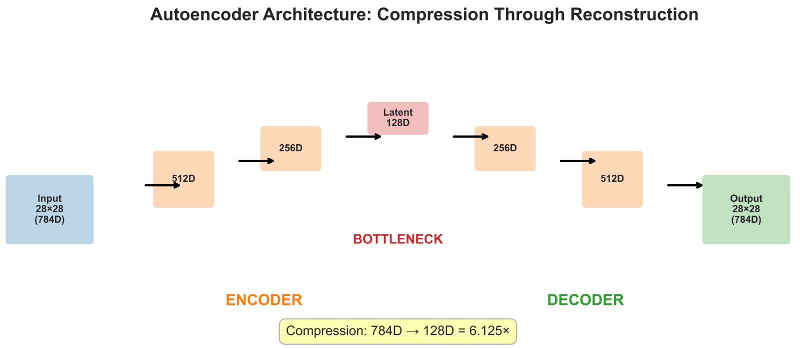
FID Score

- Range: 0-500+
- Lower = better
- Feature distance
- Real vs generated

Perplexity (Text)

- Range: 1-10,000+
- Lower = better
- Predictability
- Language fluency

Quantitative evaluation: IS=300+ (excellent), FID¡10 (photorealistic), Perplexity¡20 (human-like text)



Encoder

- 784D -¿ 128D
- $z = f_{enc}(x)$

Latent Space

- Bottleneck: 128D
- Key features

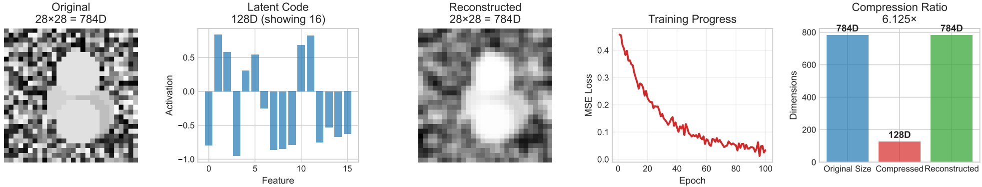
Decoder

- 128D -¿ 784D
- $\hat{x} = f_{dec}(z)$

Force information through bottleneck, learn to reconstruct

Worked Example: MNIST Compression

From 784 Pixels to 128 Features



Architecture Details:

- Input: 28x28 = 784 pixels
- Encoder: 784 -¿ 512 -¿ 256 -¿ 128
- Decoder: 128 -¿ 256 -¿ 512 -¿ 784
- Activation: ReLU (hidden), Sigmoid (output)

Training Process:

- Loss: $L = ||x - \hat{x}||^2$
- Optimizer: Adam, lr=0.001
- Epochs: 100
- Compression ratio: 784/128 = 6.125x

Reconstruction loss: MSE drops from 0.45 to 0.03 over 100 epochs

Autoencoder Successes

What Works Well

Autoencoder Successes

Visualization Placeholder

(Chart 12)

[+] SUCCESSES:

- Dimensionality reduction: 784D -¿ 128D

Quantitative Results:

- MSE: 0.031, Compression: 6.125x

Autoencoder Failures

Visualization Placeholder

(Chart 13)

[-] FAILURES:

- Blurry outputs (averaging)

Generation Metrics:

Metric	Score
IS	2.1

Averaging Problem

Visualization Placeholder

(Chart 14)

The Averaging Problem:

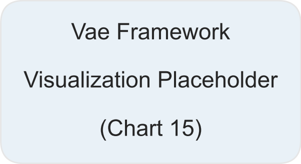
- Loss: $L = ||x - \hat{x}||^2$

Mathematical Insight:

- $\hat{x} = \arg\min E[||x - \hat{x}||^2]$

Variational Autoencoders (VAEs)

The Probabilistic Solution



Key Innovation:

- Encode to distribution, not point

VAE Loss:

$$\mathcal{L} = -E[\log p_\theta(x|z)] + KL(q||p)$$

Human Learning Analogy

How Artists Develop Mastery

Artist Learning Process

Visualization Placeholder

(Chart 16)

Traditional Art Education:

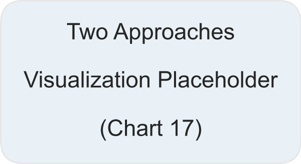
- Student creates artwork

Key Insights:

- Adversarial feedback drives improvement

Two Revolutionary Approaches

Beyond VAEs to Better Generation



Approach 1: Adversarial

- Two networks compete

Approach 2: Diffusion

- Iterative denoising

Forger Detective Analogy

Visualization Placeholder

(Chart 18)

Forger (Generator):

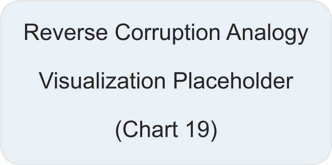
- Creates fakes from noise

Detective (Discriminator):

- Examines: real or fake?

Diffusion: The Reverse Corruption Process

Denoising in Plain English



Forward (Corruption):

- Clean image -¿ pure noise

Reverse (Generation):

- Pure noise -¿ clean image

GAN Dynamics: Geometric View

Understanding the Adversarial Process

Gan Geometric Dynamics

Visualization Placeholder

(Chart 20)

GAN Training: Step-by-Step Example

Real Loss Values from MNIST Training

Gan Training Walkthrough

Visualization Placeholder

(Chart 21)

Diffusion Mathematical Framework

Forward and Reverse Processes

Diffusion Mathematics

Visualization Placeholder

(Chart 22)

Latent Space Interpolation

Smooth Transitions in Generated Content

Latent Interpolation

Visualization Placeholder

(Chart 23)

Diffusion Denoising Visualization

From Noise to Image in 1000 Steps

Denoising Steps

Visualization Placeholder

(Chart 24)

Why Adversarial Training Works

The Mathematical Guarantee

Adversarial Theory

Visualization Placeholder

(Chart 25)

Experimental Validation

Quality Metrics Throughout Training

Quality Metrics Over Time

Visualization Placeholder

(Chart 26)

Implementation: Stable Diffusion API

Production-Ready Generative AI

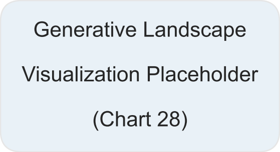
Stable Diffusion Api

Visualization Placeholder

(Chart 27)

The Generative AI Landscape

Four Fundamental Approaches



VAEs: Probabilistic, smooth latent, blurry
GANs: Adversarial, sharp outputs, unstable

Diffusion: Iterative denoising, high quality, slow
Transformers: Sequential, excellent text, scalable

Each approach has unique strengths - modern systems combine techniques

Fundamental Trade-offs

No Free Lunch in Generative Modeling

Generative Tradeoffs

Visualization Placeholder

(Chart 29)

Training Stability:

- VAEs, Diffusion: Stable
- GANs: Unstable

Quality:

- Diffusion, GANs: Excellent
- VAEs: Blurry

State-of-the-Art Applications

Production Generative AI Systems

Modern Applications

Visualization Placeholder

(Chart 30)

Image Generation:

- DALL-E 3, Midjourney
- Stable Diffusion, Firefly

Text Generation:

- GPT-4, Claude, Gemini
- Llama 2 (open)

Summary & Ethical Considerations

Power and Responsibility in Generative AI

Ethics Summary

Visualization Placeholder

(Chart 31)

Capabilities:

- Realistic images from text
- Human-like writing

Challenges:

- Deepfakes, misinformation
- Copyright issues