

Week 0e: Generative AI

The Creation Challenge

Machine Learning for Smarter Innovation

BSc-Level Course

October 6, 2025

- 1 Act 1: The Challenge
- 2 Act 2: Variational Autoencoders
- 3 Act 3: Adversarial & Diffusion
- 4 Act 4: Synthesis

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

Discriminative: “Is this a cat?” — Generative: “Draw me a cat”

Generative Models

Learn: $P(x)$ or $P(x, y)$

“What does the data distribution look like?”

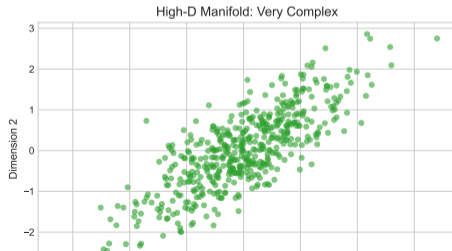
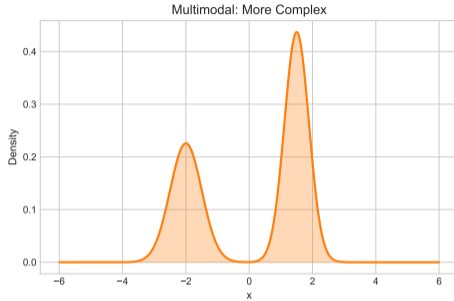
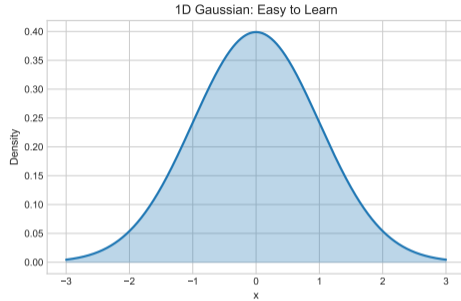
Examples:

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

Goal: Data generation

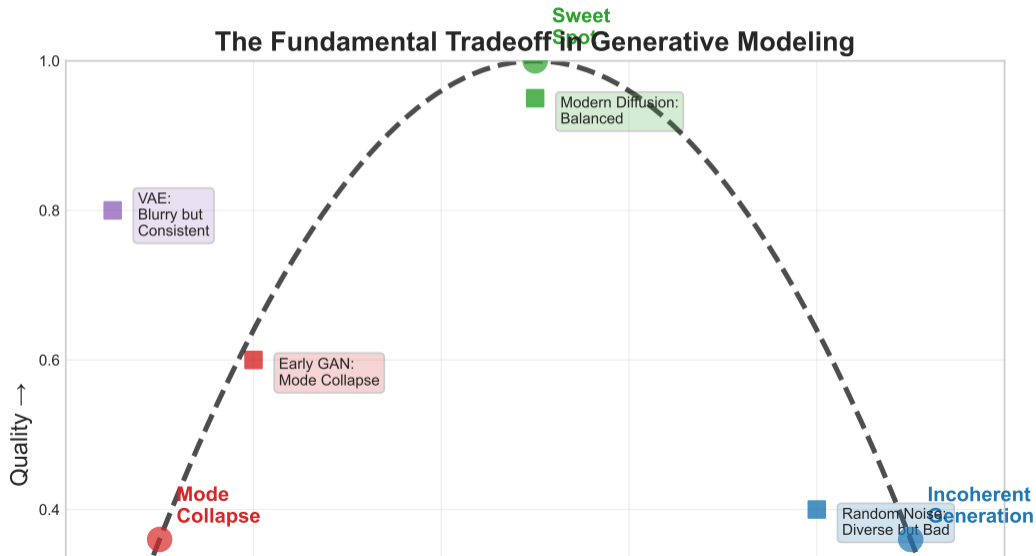
The Hard Problem

Why Generation is Fundamentally Difficult



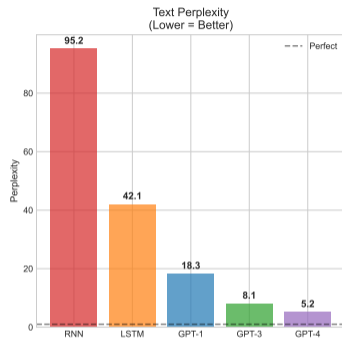
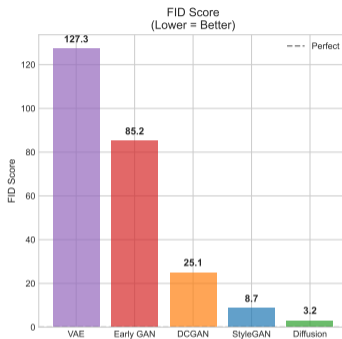
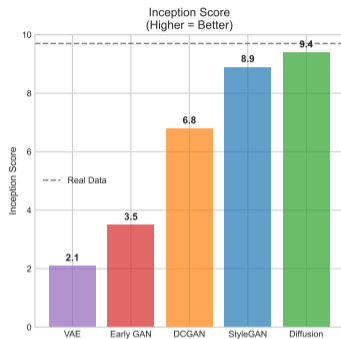
The Fundamental Tradeoff

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)

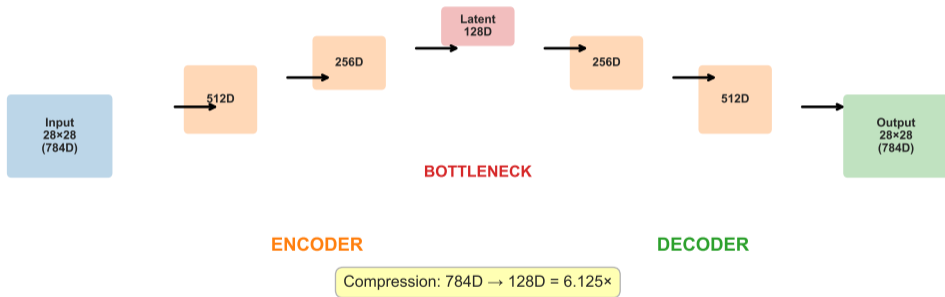
- Range: 1-10,000+
- Lower = better
- Predictability
- Language fluency

Quantitative evaluation: IS=300+ (excellent), FID_i10 (photorealistic), Perplexity_i20 (human-like text)

Autoencoders: The Foundation

Learning Compressed Representations

Autoencoder Architecture: Compression Through Reconstruction



Encoder

Latent Space

Decoder

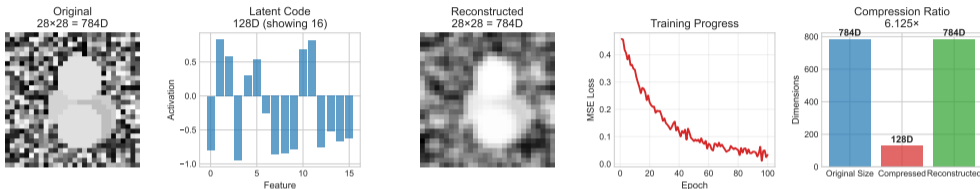
- Input: 784D (28x28 image)

- Bottleneck: 128D

- Latent: 128D

Worked Example: MNIST Compression

From 784 Pixels to 128 Features



Architecture Details:

- Input: $28 \times 28 = 784$ pixels
- Encoder: $784 \rightarrow 512 \rightarrow 256 \rightarrow 128$
- Decoder: $128 \rightarrow 256 \rightarrow 512 \rightarrow 784$
- Activation: ReLU (hidden), Sigmoid (output)

Training Process:

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

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

Autoencoder Successes
Visualization Placeholder
(Chart 12)

Autoencoder Failures
Visualization Placeholder
(Chart 13)

Root Cause Analysis

Why Autoencoders Generate Poorly

Averaging Problem
Visualization Placeholder
(Chart 14)

Vae Framework
Visualization Placeholder
(Chart 15)

Artist Learning Process
Visualization Placeholder
(Chart 16)

Two Revolutionary Approaches

Beyond VAEs to Better Generation

Two Approaches
Visualization Placeholder
(Chart 17)

GANs: The Forger vs Detective Game

Adversarial Training in Plain English

Forger Detective Analogy
Visualization Placeholder
(Chart 18)

Diffusion: The Reverse Corruption Process

Denoising in Plain English

Reverse Corruption Analogy
Visualization Placeholder
(Chart 19)

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 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)

Adversarial Theory
Visualization Placeholder
(Chart 25)

Quality Metrics Over Time
Visualization Placeholder
(Chart 26)

Implementation: Stable Diffusion API

Production-Ready Generative AI

Stable Diffusion Api
Visualization Placeholder
(Chart 27)

Generative Landscape
Visualization Placeholder
(Chart 28)

Generative Tradeoffs
Visualization Placeholder
(Chart 29)

Modern Applications
Visualization Placeholder
(Chart 30)

Summary & Ethical Considerations

Power and Responsibility in Generative AI

Ethics Summary
Visualization Placeholder
(Chart 31)