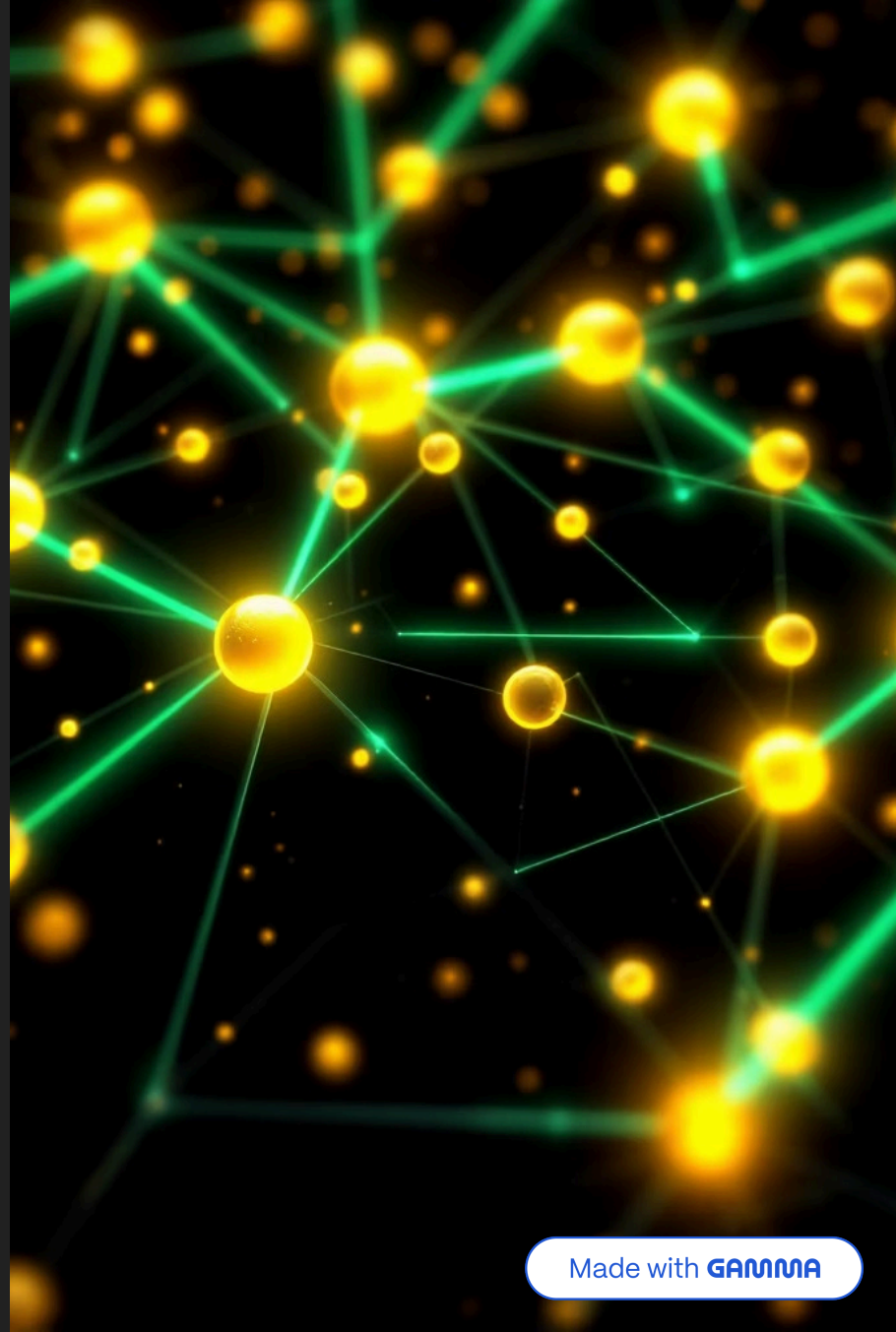
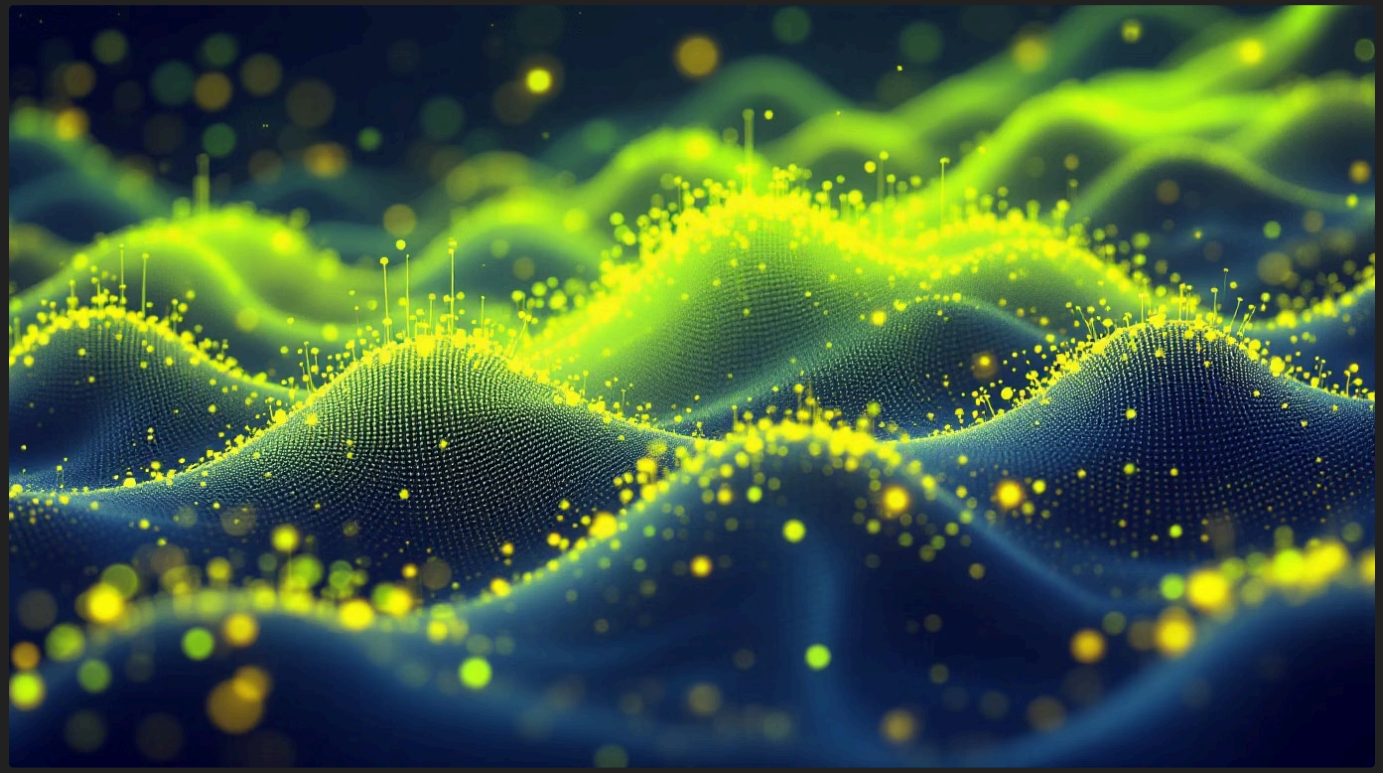
Common AI and ML Algorithms

A Technical Overview for Industry Professionals



Made with GAMMA

K-Means Clustering



Unsupervised Learning

Iteratively assigns data points to the nearest cluster centroid, grouping similar observations without labeled guidance.

Key Parameters

- Number of clusters (k)
- Distance metric selection
- Initialization method

Applications

Customer segmentation, anomaly detection, image compression, and market basket analysis.

Made with GAMMA

Linear Regression

Mathematical Model

$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n + \varepsilon$

Models continuous outcomes as linear combination of input variables with error term.

Core Assumptions

- Linearity and independence
- Homoscedasticity
- Normal error distribution

Real-World Use

Forecasting sales, risk modeling, trend analysis, and predictive analytics across industries.



Made with GAMMA

Random Forests

Ensemble Method

Combines multiple decision trees using bagging (bootstrap aggregating) to improve accuracy and robustness.

Reduced Overfitting

Random feature selection at each split prevents individual trees from memorizing training data.

Missing Data

Handles missing values gracefully through surrogate splits and imputation during training.

Applications: Fraud detection in financial transactions, recommendation systems, bioinformatics, and healthcare diagnostics.



Made with GAMMA

Gradient Boosting

Sequential Optimization

Builds models incrementally where each new tree corrects errors of previous learners. Unlike random forests, trees are dependent and ordered.

Popular Frameworks

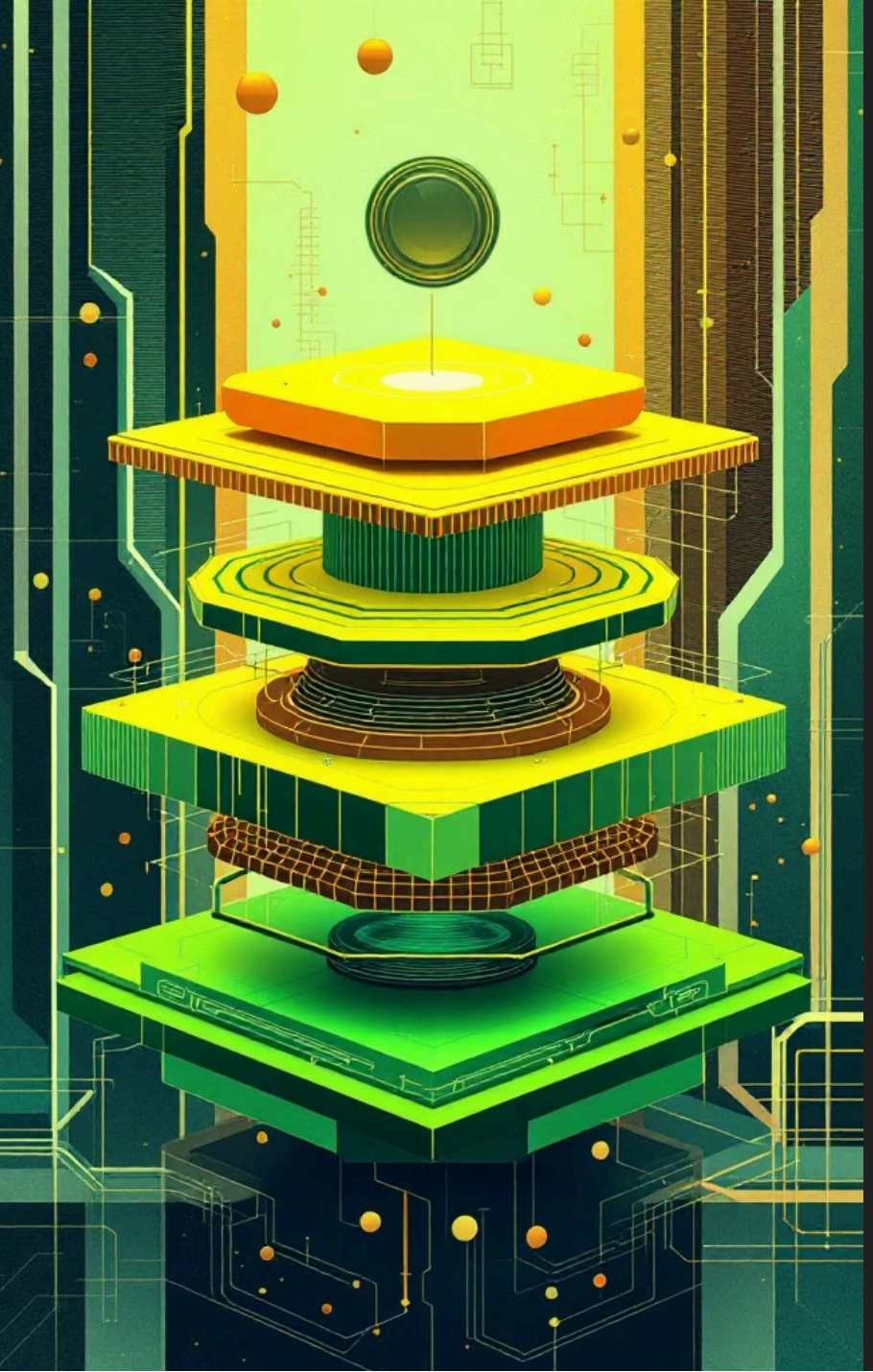
- **XGBoost:** Extreme gradient boosting with regularization
- **LightGBM:** Gradient-based one-side sampling for speed
- **CatBoost:** Automatic categorical feature handling

Use Cases

Credit scoring, search ranking, predictive maintenance, and high-stakes classification problems.



Made with GAMMA



Convolutional Neural Networks

Convolutional Layers

Apply learnable filters to detect spatial patterns and features at multiple scales.

Pooling Operations

Downsample feature maps to reduce dimensionality while preserving important information.

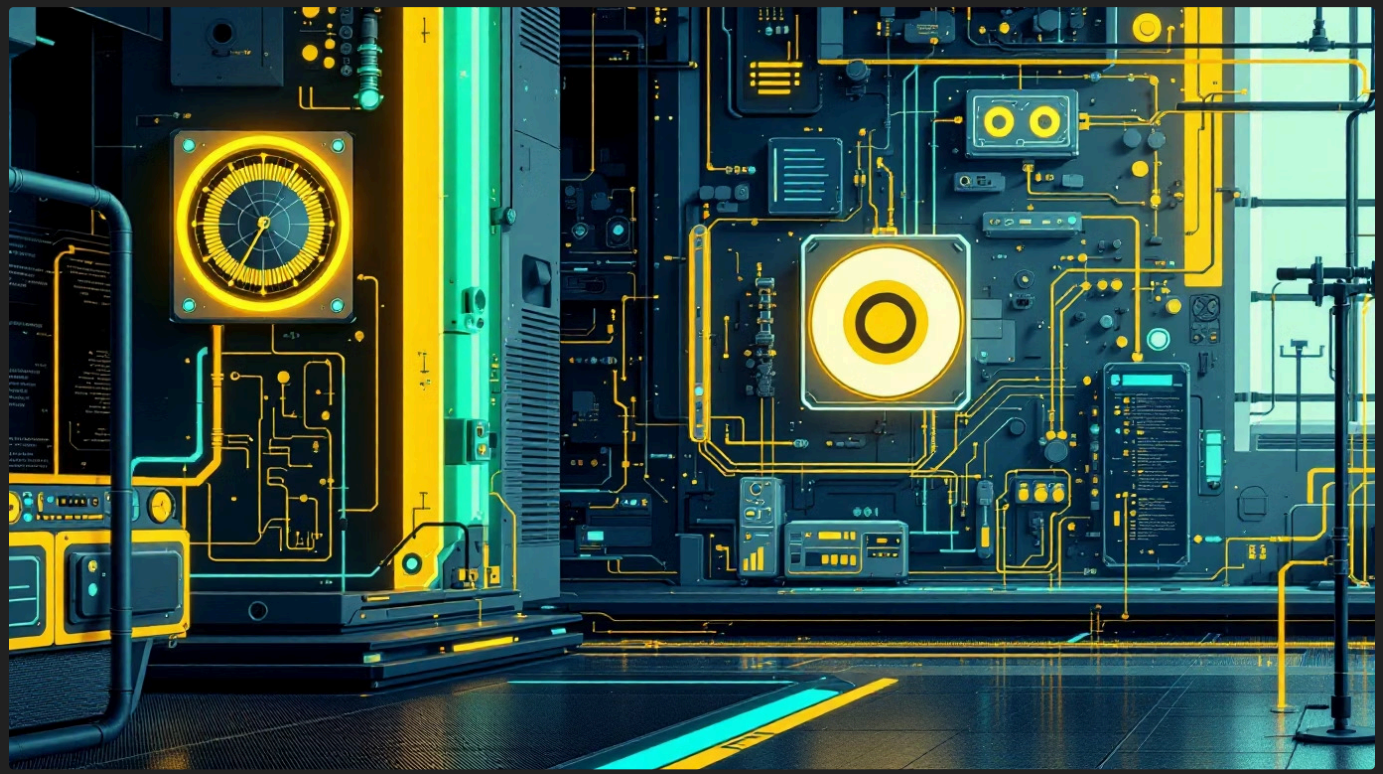
Fully Connected

Final layers interpret extracted features for classification or regression tasks.

Applications: Image recognition (ResNet, VGG), medical imaging diagnostics, autonomous vehicle perception, and satellite image analysis.

Made with GAMMA

GPT: Generative Pre-trained Transformer



Architecture Foundation

Transformer-based architecture leveraging self-attention mechanisms to weigh input importance dynamically across sequence positions.

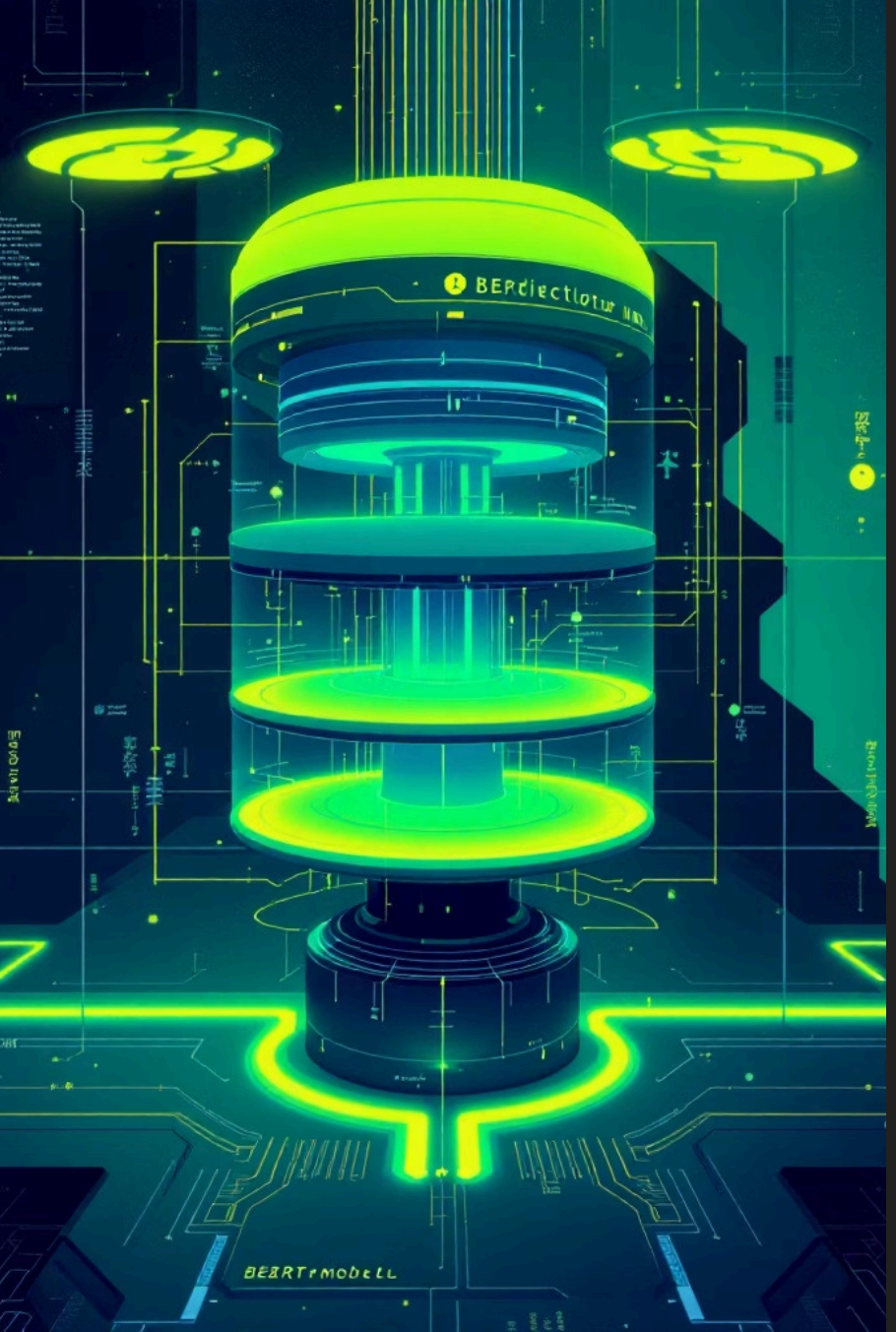
Training Paradigm

Pre-trained on massive text corpora using unsupervised learning, then fine-tuned for specific downstream tasks with labeled data.

Applications

- Chatbots and conversational agents
- Text summarization and translation
- Code generation and completion
- Creative content creation

Made with GAMMA



BERT: Bidirectional Encoder Representations

1

Bidirectional Training

Context from both left and right directions enables deeper understanding of word meaning in context.

2

Premise:

Context determines meaning—same word, different usage.

3

Fine-Tuning

Adaptable architecture for diverse NLP tasks: question answering, sentiment analysis, named entity recognition.

Applications: Search optimization, NLP pipelines, document classification, and semantic similarity tasks.

Made with GAMMA

Generative Adversarial Networks

Adversarial Architecture

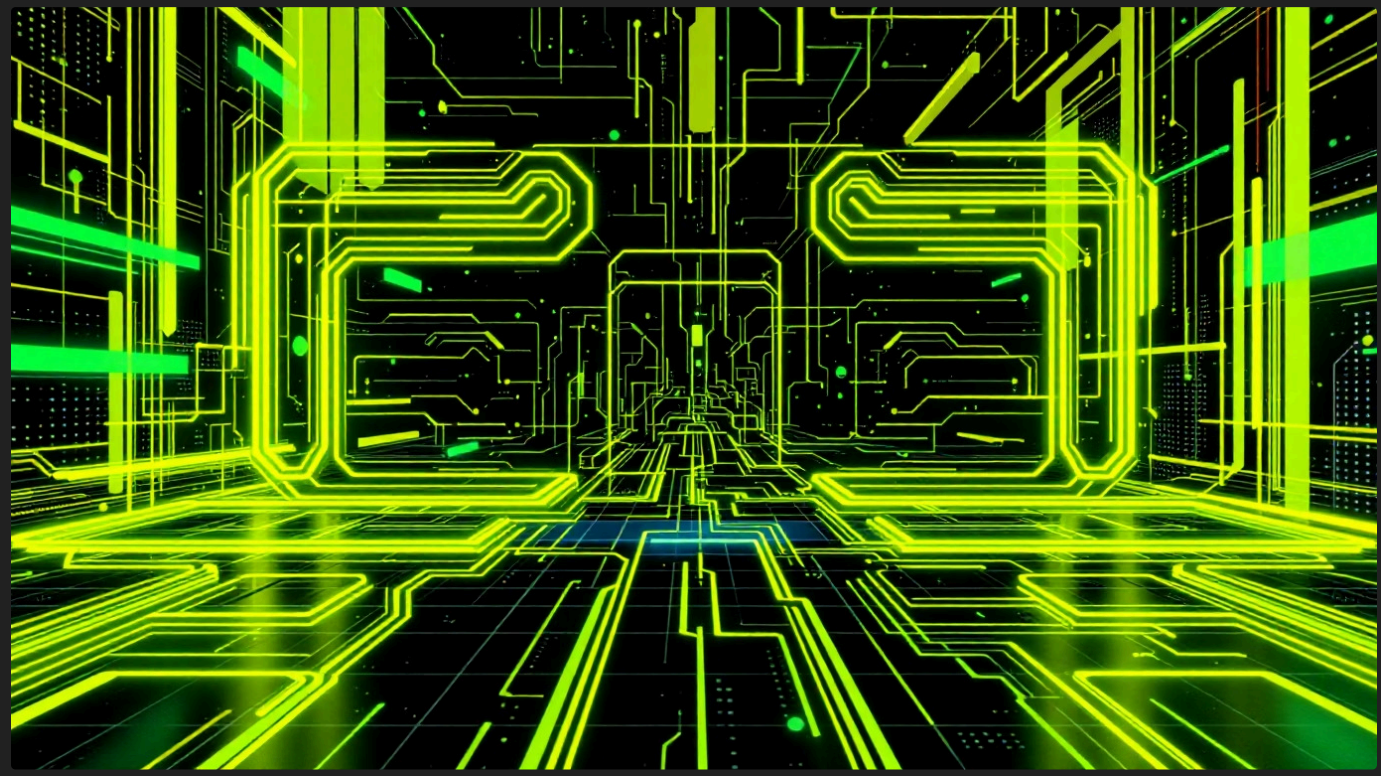
Generator: Creates synthetic data from random noise

Discriminator: Distinguishes real from fake samples

Two networks compete in minimax game until generator produces realistic data.

Challenges

- Mode collapse (limited diversity)
- Training instability
- Hyperparameter sensitivity



Applications

Image synthesis (StyleGAN), data augmentation for ML, deepfake detection, and artistic generation.

Made with GAMMA



Key Takeaways

Diverse Methods

AI and ML algorithms span supervised, unsupervised, and deep learning paradigms with distinct strengths.

Ensemble Dominance

Random forests and gradient boosting remain industry standards for tabular data problems.

Deep Learning Era

Transformer architectures and CNNs power modern computer vision and NLP applications.

Continuous innovation drives model efficiency, interpretability, and accessibility across industries.

Made with GAMMA