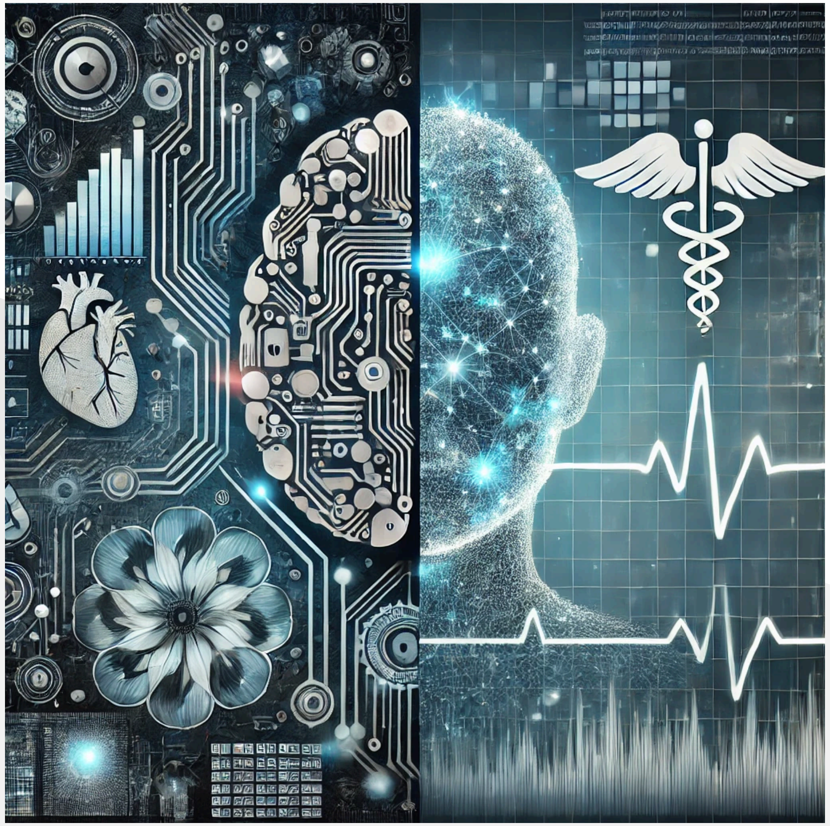


Deep Learning

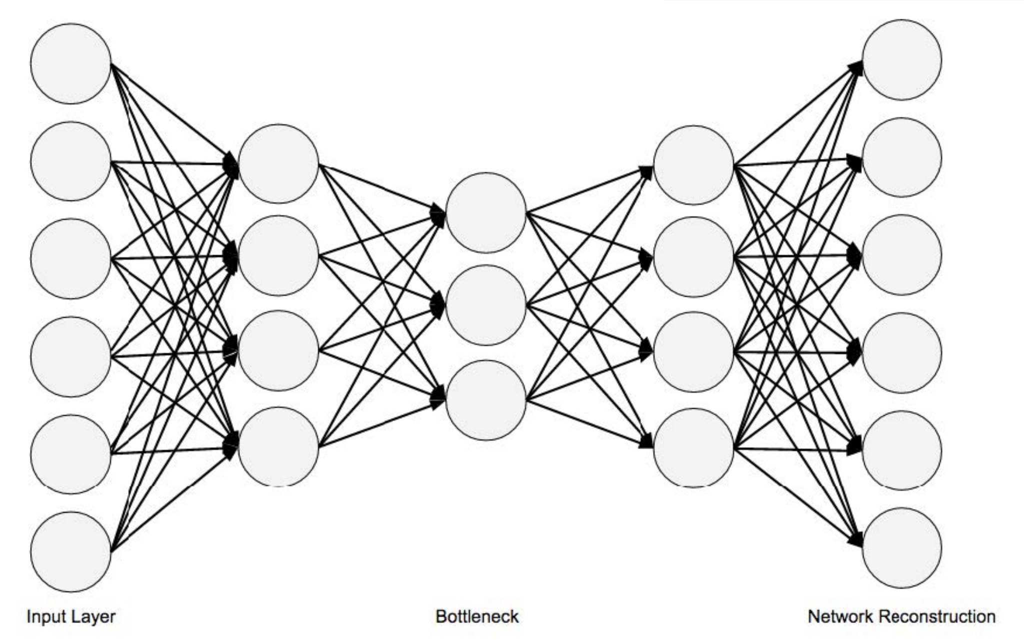
Major Architectures of Deep Networks

Introduction to Network Architectures

- To generate data (e.g., images, audio, or text):
  - GANs
  - VAEs
  - Recurrent Neural Networks
- To model images:
  - CNNs
  - DBNs
- To model sequence data:
  - Recurrent Neural Networks/LSTMs

Autoencoders

Autoencoders learn to encode data into a smaller dimensionality (latent space) and then decode it back to its original form.

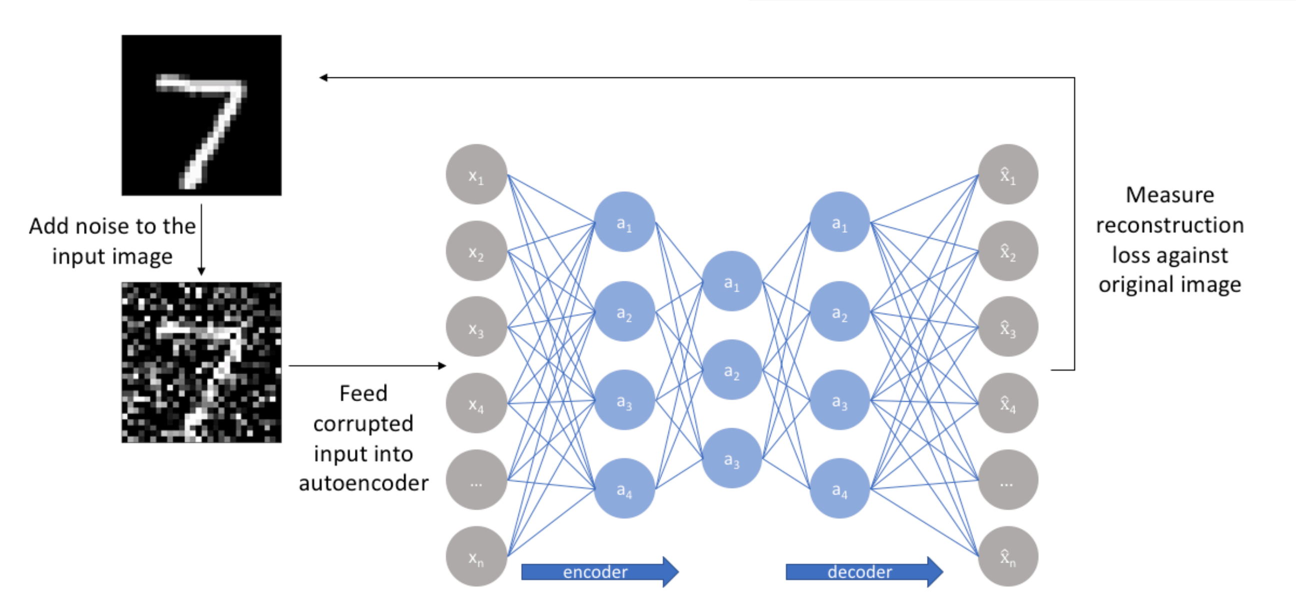


Autoencoders

- Comprised of two main components:

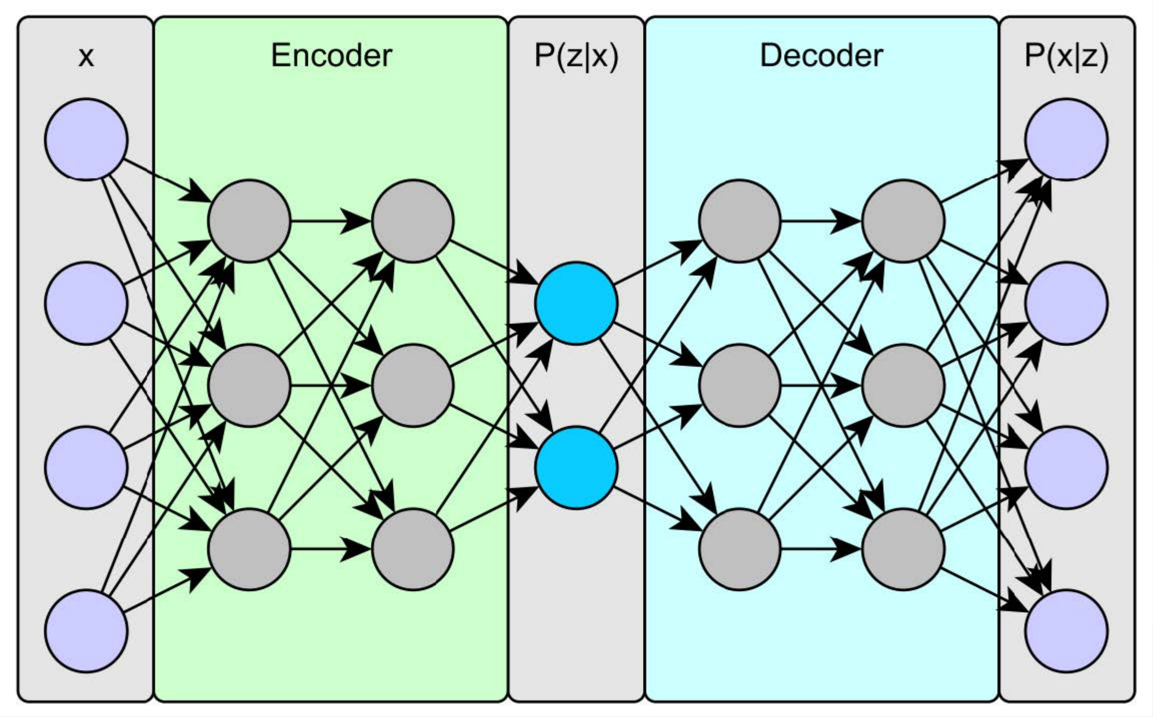
  Encoder: Compresses input into a lower-dimensional representation.

  Decoder: Reconstructs the original input from the encoded representation.



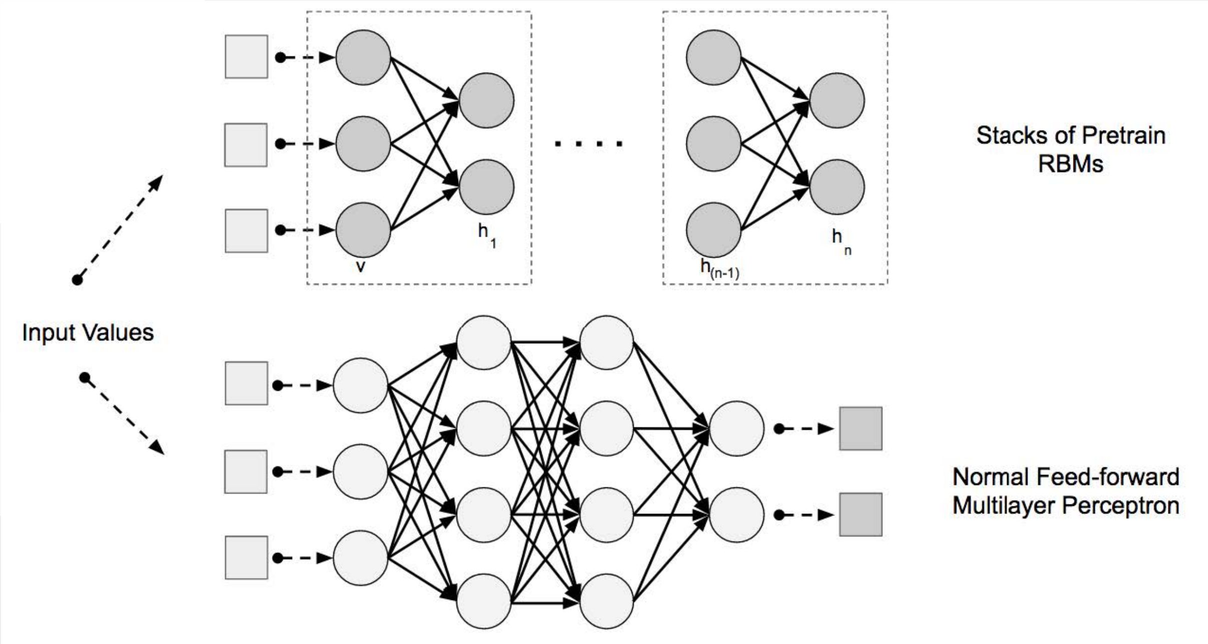
Variational Autoencoder (VAE)

Adds stochastic properties for more robust feature extraction.



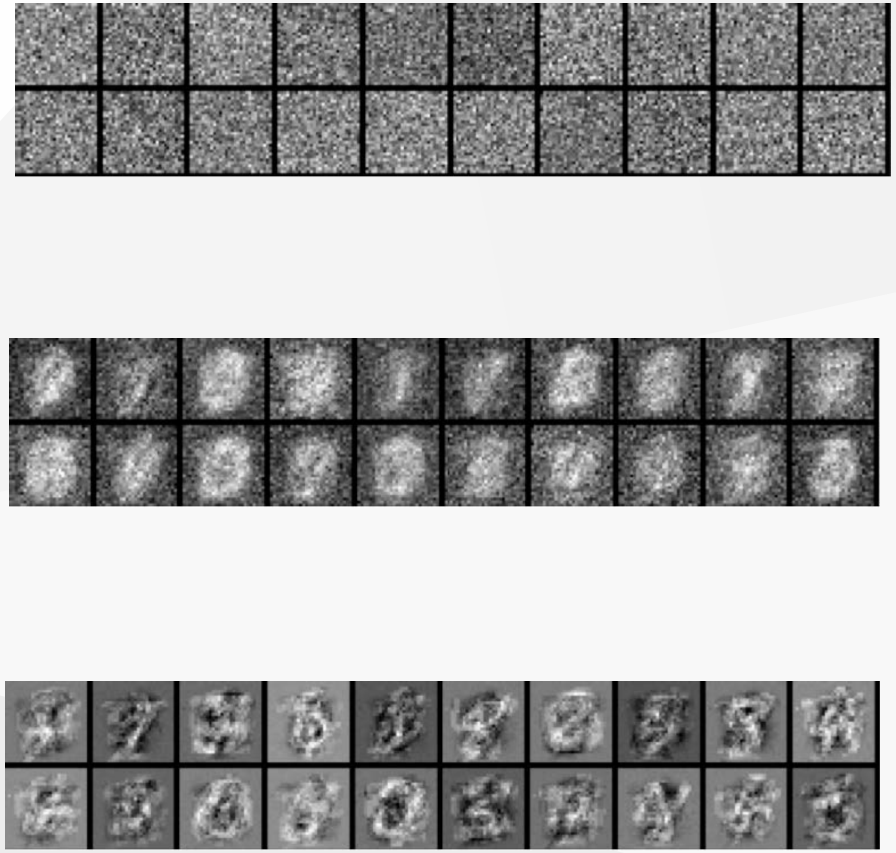
Deep Belief Networks (DBNs)

- **Components**:
    - Stacked Restricted Boltzmann Machines (RBMs).
    - Fine-tuning phase with supervised learning.



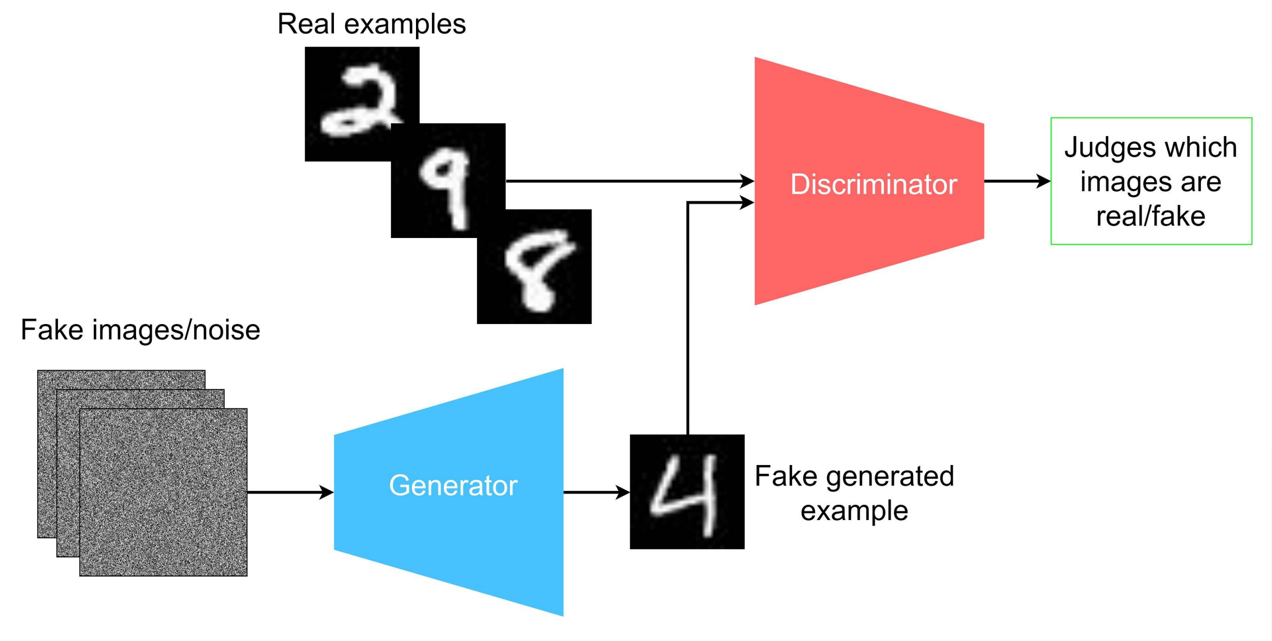
Deep Belief Networks (DBNs)

- **Purpose**:
  - DBNs use unsupervised layers to capture high-level data features before applying a classifier.



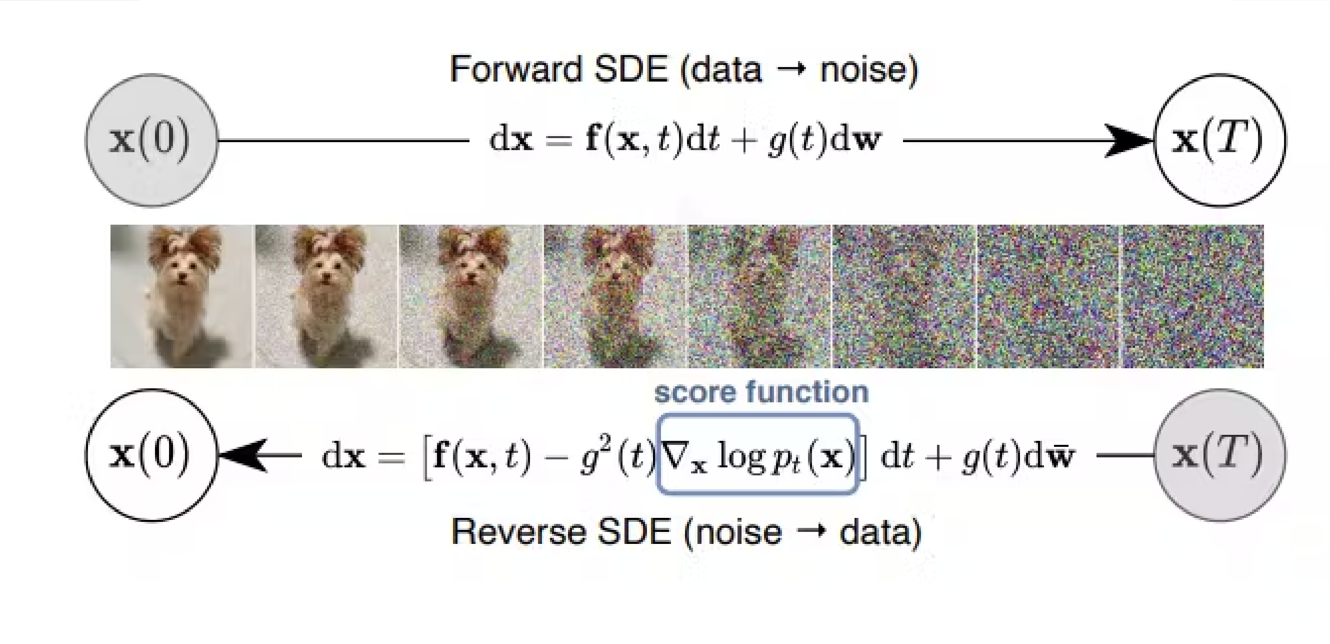
Generative Adversarial Networks (GANs)

They are composed of a **Generator** and **Discriminator** network. During training the generator tries to produce realistic data while the discriminator distinguishes between real and fake data..



Diffusion Models

- **Definition**: Probabilistic models that generate data by reversing a diffusion process.



Starts with Gaussian noise and gradually refines it to generate samples.

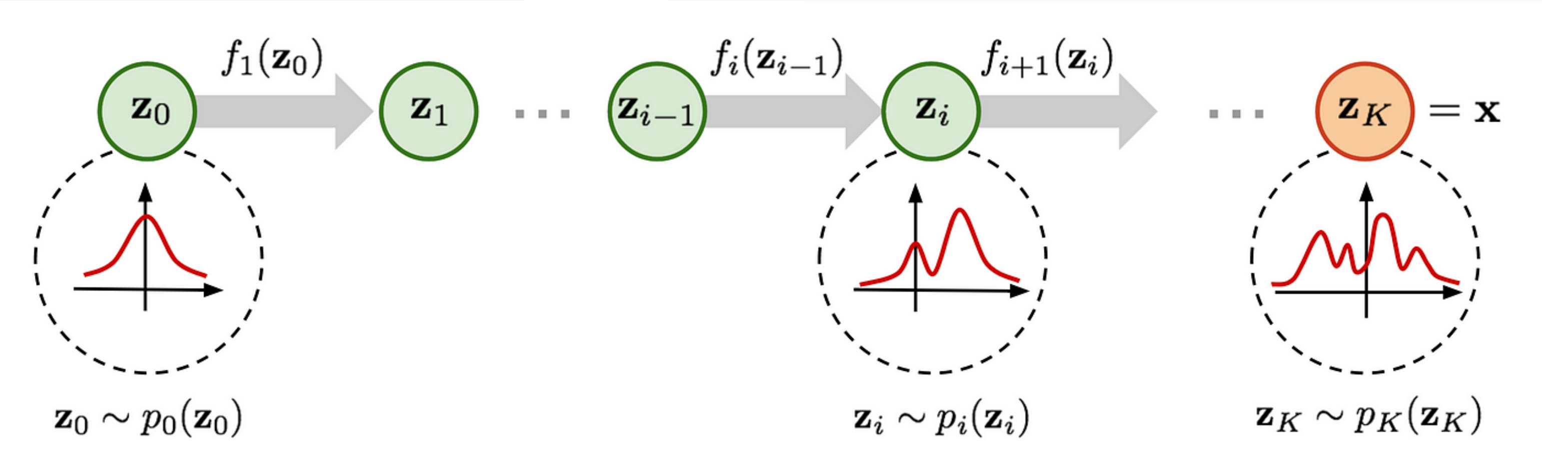
Trained by simulating and then reversing a gradual noising process.

Normalizing Flows

- **Definition**: Models that learn complex data distributions by transforming simple probability distributions through a series of invertible mappings.

Bijective Transformation: Ensures data can be transformed back to original space.

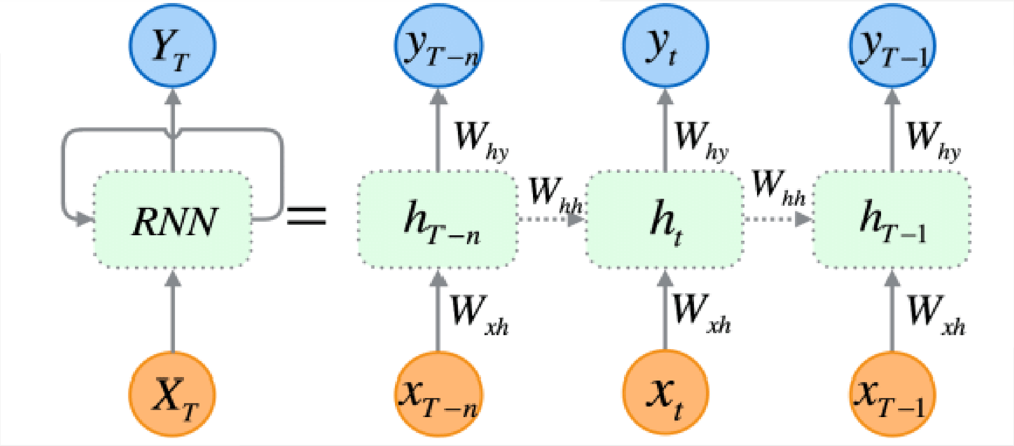
Likelihood Maximization: Direct computation of probability densities.



Recurrent Neural Networks (RNNs)

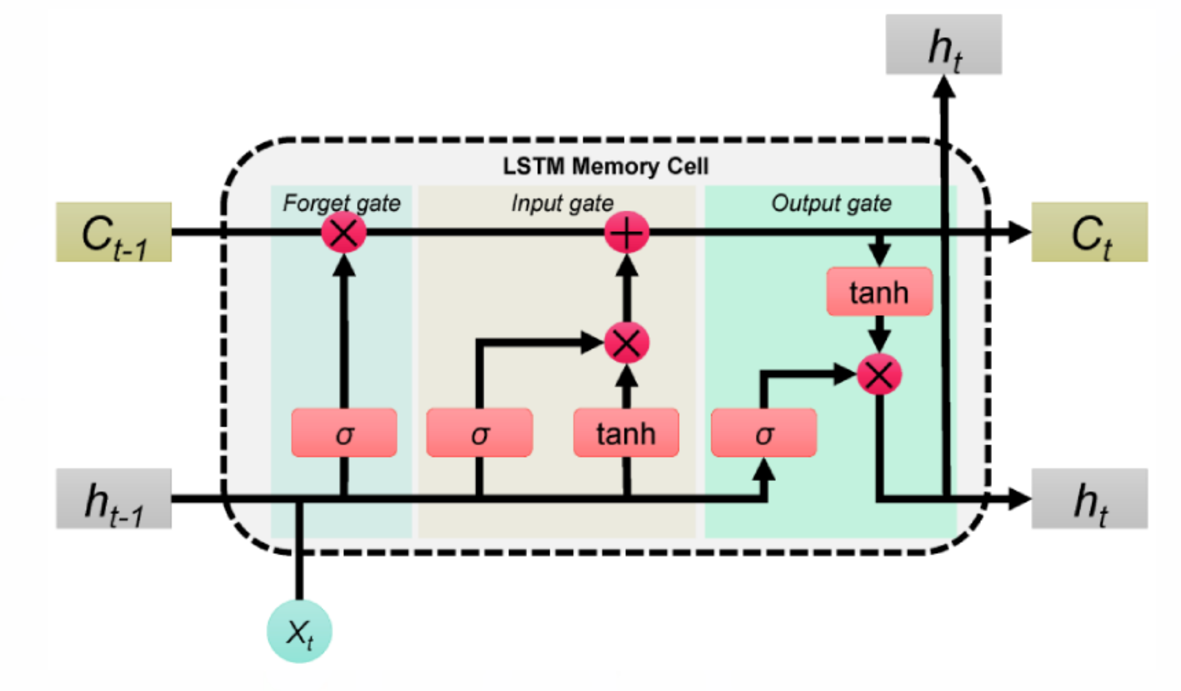
They model sequential data where time dependency matters.

Hidden State: Maintains context across time steps.



Long Short-Term Memory (LSTM) Networks

LSTM networks are the most common variation of RNNs. The critical components are the **memory cell** and the **gates**. The gating structure allows information to be retained across many time-steps. This allows gradients to flow across many time-steps, overcoming the **vanishing gradient** problem that occurs with most RNNs.

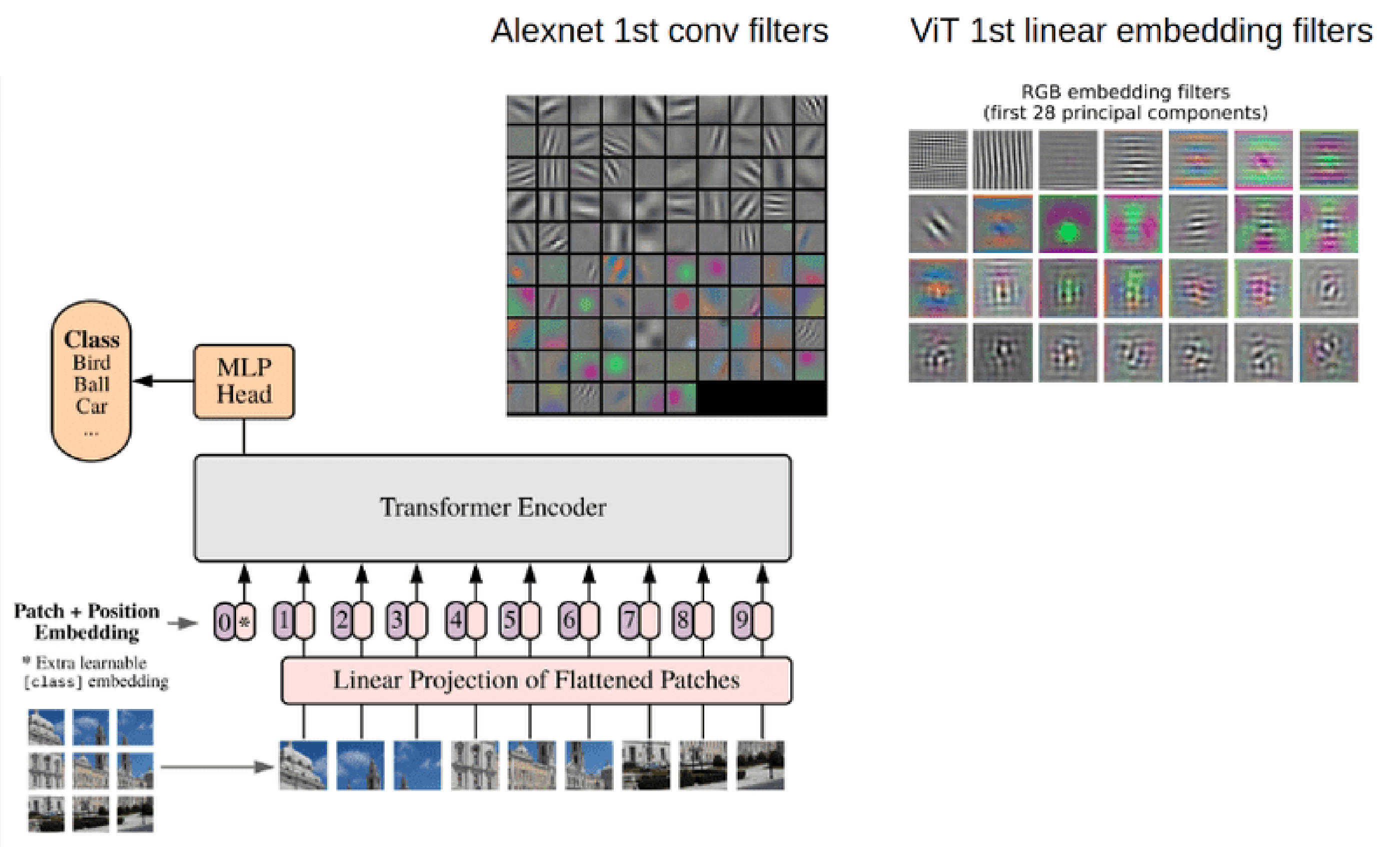


Recursive Neural Networks

- **Definition**: Networks with a tree-like structure to capture hierarchical data.
- **Applications**:
  - Natural Language Processing (NLP) and image scene decomposition.
  - In biomedical contexts, used for parsing complex structures in genomic sequences or biological images.
- **Advantages**:
  - Handles structured data and hierarchical relationships .

Transformers

Core Mechanism: **Self-Attention** allows each data point to attend to all others, capturing relationships regardless of their position.



Advantage:

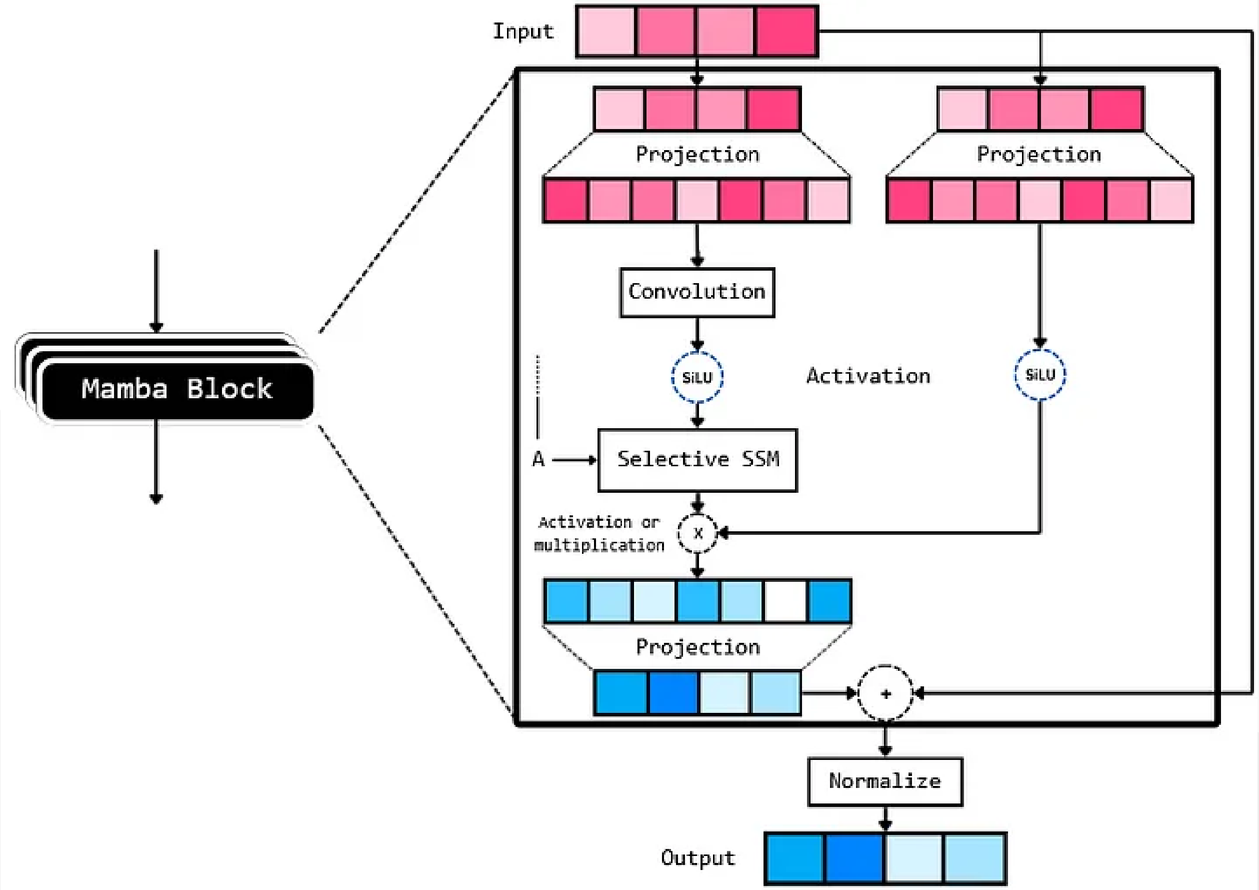
No sequential constraints like RNNs; can handle long-range dependencies more effectively.

Mamba Models

Mamba models are built upon Structured State Space Models (SSMs), which represent dynamic systems through state variables.

Mamba enhances SSMs by introducing mechanisms to dynamically adjust their focus based on input data, enabling more efficient handling of complex, long-sequence tasks.

While Transformers scale by $O(n^2)$, Mamba scales by $O(n)$!



Convolutional Neural Networks (CNNs)

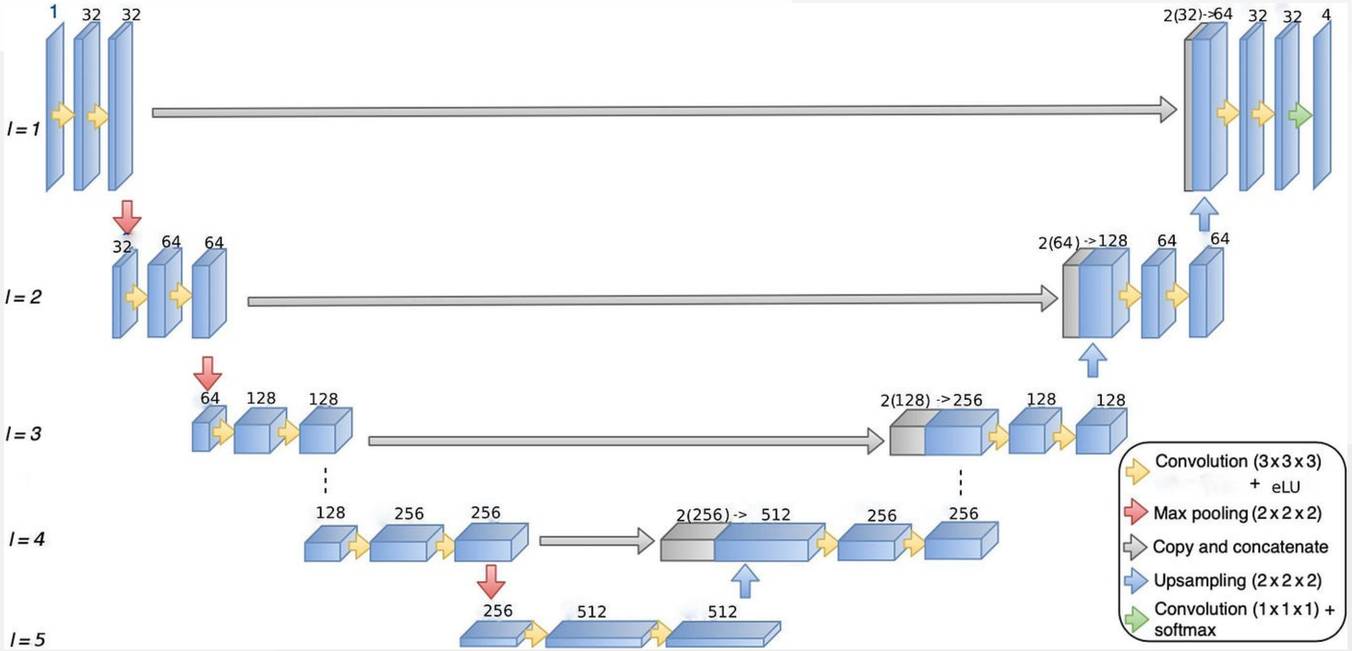
They are designed primarily for image processing tasks.

Core Components:

- **Convolution Layers**: Extract spatial features.
- **Pooling Layers**: Downsample feature maps to reduce dimensionality and control overfitting.
- **Fully Connected Layers**: Often used for final classification .

UNet

The U-Net is a convolutional neural network architecture that was designed for fast and precise segmentation of images. It has performed extremely well in several challenges and to this day, its application has spread also to other fields such as image denoising.

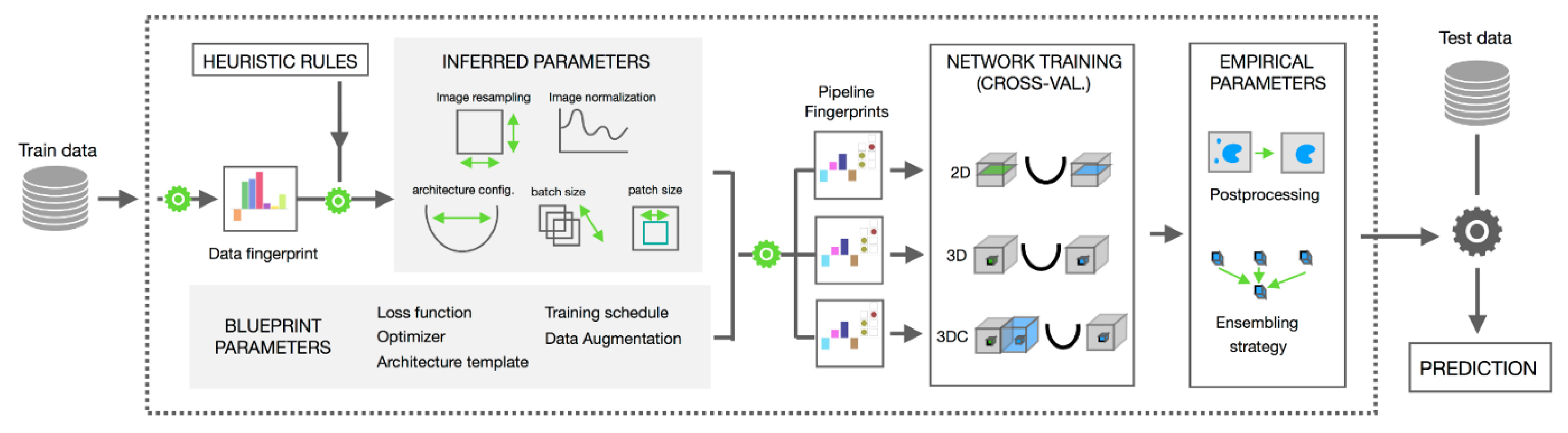


nnUNet

nnUNet is an out-of-the-box 3D medical image segmentation framework that automates model configuration and optimization.

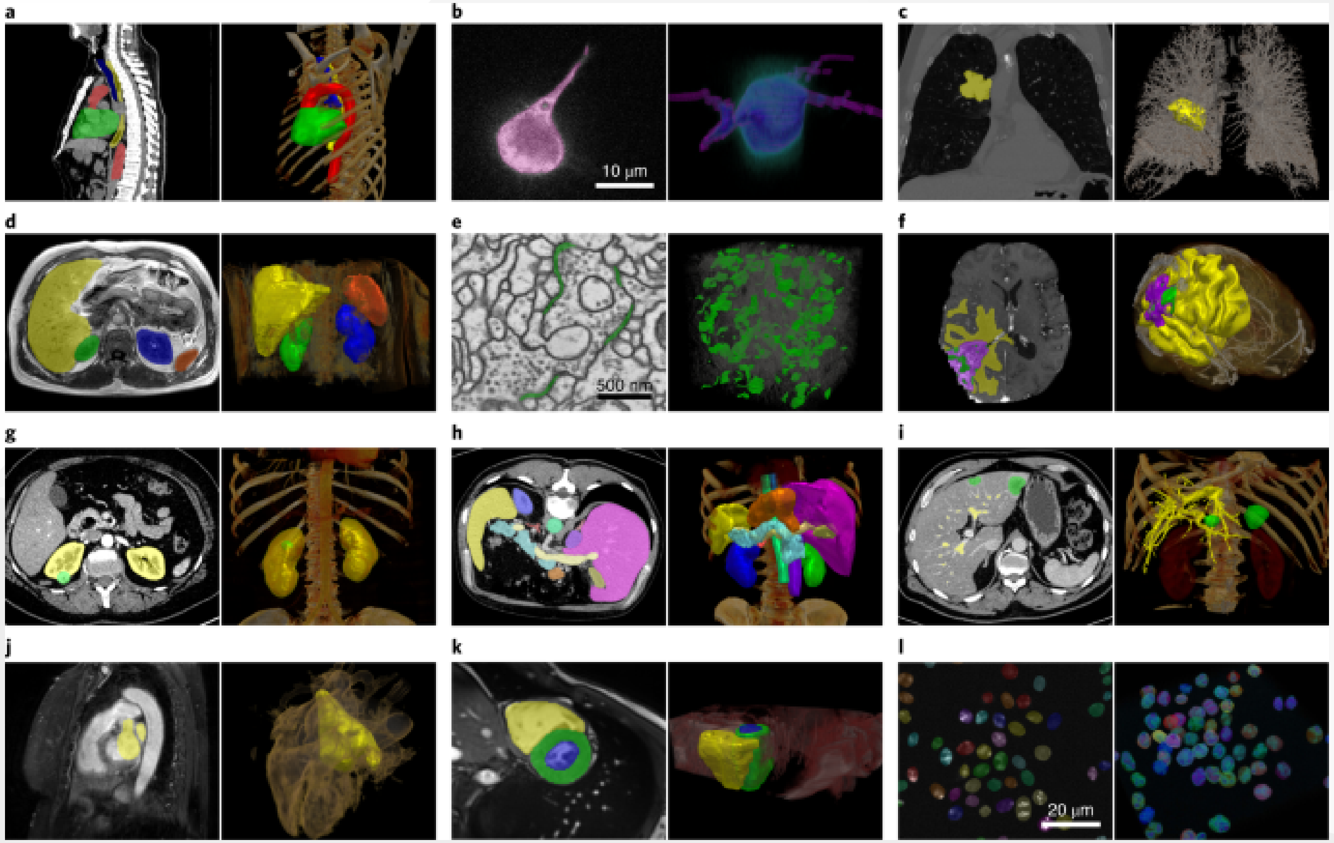
Self-Configuration: Adapts network architecture, hyperparameters, and training schemes to the dataset.

Ensemble Learning: Combines predictions from multiple models for more robust results.



nnUNet

nnUNet was the top-performing model on the **Medical Segmentation Decathlon**, excelling across diverse datasets. It has become the standard in biomedical imaging segmentation due to its adaptability.



Future Directions

- **Model Personalization**: Tailoring deep networks to individual patient data.
- **Integration of Multi-Modal Data**: Combining imaging, genomics, and EHR data.
- **Ethics and Fairness**: Ensuring AI fairness and reducing biases in healthcare applications.
- **Emerging Architectures**:
    - Hybrid models combining CNNs, RNNs, and attention mechanisms for richer modeling.