

Introduction to Deep Learning

Introduction to Deep Learning

An overview of the fundamental concepts, architecture, and practical considerations.

1. What Is Deep Learning?

Aspect	Description
Definition	A sub-field of machine learning that uses neural networks with many layers ("deep" networks) to automatically learn hierarchical representations from data.
Goal	Capture complex, non-linear relationships without manual feature engineering.
Typical Data Types	Images, audio, video, text, time-series, and multimodal signals.
Why "deep"?	Stacking multiple layers allows the model to transform raw inputs into increasingly abstract features (e.g., edges → shapes → objects).

2. Core Building Blocks

Component	Role	Common Variants
Neuron (Unit)	Computes a weighted sum of its inputs, adds a bias, and passes the result through an activation function.	Linear, ReLU, Leaky-ReLU, ELU, Sigmoid, Tanh
Layer	A collection of neurons that operate in parallel on the same input.	Fully-connected (dense), Convolutional, Recurrent, Pooling, Normalization
Network (Model)	An ordered stack of layers that defines a mapping from input to output.	Feed-forward (CNN, MLP), Recurrent (RNN, LSTM, GRU), Transformer, Autoencoder, GAN
Loss Function	Quantifies the discrepancy between predictions and ground-truth labels.	Cross-entropy, MSE, Hinge, KL-divergence, CTCLoss
Optimizer	Updates model parameters to minimize the loss.	SGD, Momentum, Adam, RMSprop, AdamW

Introduction to Deep Learning

Component	Role	Common Variants
Regularization	Prevents overfitting and improves generalization.	L1/L2 weight decay, Dropout, BatchNorm, Data augmentation, Early stopping

3. Popular Network Architectures

Architecture	Primary Use Case	Key Idea
Multilayer Perceptron (MLP)	Tabular data, simple classification/regression	Stacked dense layers with non-linear activations
Convolutional Neural Network (CNN)	Images, video, spatial data	Convolution + pooling to exploit local spatial correlation
Recurrent Neural Network (RNN)	Sequential data (text, speech)	Hidden state carries information across time steps
Long Short-Term Memory (LSTM) / GRU	Long-range dependencies in sequences	Gated mechanisms mitigate vanishing gradients
Transformer	NLP, vision language, speech, protein folding	Self-attention replaces recurrence; parallelizable
Autoencoder (AE)	Dimensionality reduction, denoising, pre-training	Encoder compresses, decoder reconstructs input
Generative Adversarial Network (GAN)	Image synthesis, style transfer	Two networks (generator vs. discriminator) compete in a minimax game

4. Training Deep Networks

1. Data Preparation

- Clean, normalize, and optionally augment data.
- Split into training / validation / test sets (typical ratios: 70/15/15).

2. Forward Pass

Introduction to Deep Learning

- Input ? layers ? output predictions.

3. Loss Computation

- Compare predictions to targets using an appropriate loss.

4. Backward Pass (Back?Propagation)

- Compute gradients of loss w.r.t. each parameter via the chain rule.
- Use automatic differentiation frameworks (PyTorch, TensorFlow).

5. **Parameter