**Unit 1  
Introduction to Deep Learning & Neural Network Foundations**

**1.1 What is Deep Learning?**

Deep Learning is a subset of Machine Learning that uses artificial neural networks (ANNs) with multiple hidden layers to automatically learn representations from raw data. Unlike traditional ML, it reduces dependence on handcrafted features by discovering hierarchical abstractions: edges → shapes → objects → concepts.  
Modern Deep Learning powers applications in computer vision (image classification, detection, medical imaging), NLP (translation, chatbots, summarization), speech recognition, recommendation systems, and robotics. Its success comes from the convergence of large datasets, GPU acceleration, and improved algorithms.

**1.2 Brief History of Neural Networks**

• **1943** – McCulloch & Pitts: Proposed the first mathematical model of a neuron.  
• **1958** – Frank Rosenblatt: Introduced the Perceptron, capable of linear classification.  
• **1969** – Minsky & Papert: Exposed perceptron limitations (e.g., failure on XOR), triggering skepticism.  
• **1970s –** AI Winter: Decline in neural network research due to limited computing and poor results.  
• **1980s** – Geoffrey Hinton & colleagues: Revived the field with backpropagation, enabling training of multi-layer networks.  
• **2000s**–present: Growth of deep architectures (CNN, RNN, LSTM, Transformers) fueled by data and compute, leading to breakthroughs in applied AI.

**1.3 The Perceptron**

The Perceptron, designed by Rosenblatt in 1958, is the foundational building block of neural networks.  
• Inputs are multiplied by corresponding weights.  
• A summation function computes the weighted sum.  
• An activation function applies a threshold to decide the output (binary classification).  
Strength: Solves linearly separable problems.  
Limitation: Cannot solve non-linear functions like XOR.  
The perceptron introduced the idea of learning through weight updates, which influenced all later models.

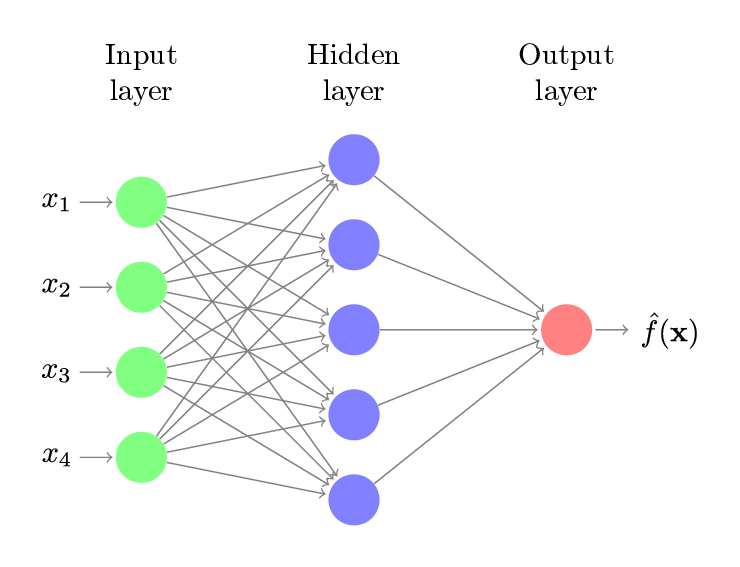
**1.4 1980 – Geoffrey Hinton & Backpropagation**

Backpropagation redefined neural network training by allowing multi-layer perceptrons to learn complex mappings.  
• **Forward Pass**: Data flows input → hidden → output.  
• **Error Calculation:** Difference between predicted and actual output.  
• **Backward Pass:** Error propagated back using the chain rule of calculus.  
• **Weight Update:** Adjustments made via gradient descent to minimize error.  
Impact: Neural networks became trainable at scale, reviving interest and research momentum in the 1980s.

**1.5 Impact of Backpropagation on ANN, CNN, RNN**

• ANN: Enabled practical training of multi-layer feedforward networks for classification/regression.  
• CNN: Allowed deeper convolutional models for vision tasks (e.g., ImageNet breakthroughs).  
• RNN: Facilitated sequential learning by adjusting recurrent connections through time.  
Backpropagation became the core algorithm powering nearly all deep learning architectures.

**1.6 Basic Neural Network Structure**



A standard neural network is organized into interconnected layers of neurons:  
• Input Layer – Accepts raw features (e.g., pixels, numerical values, embeddings).  
• Hidden Layers – Transform inputs via weighted connections to learn abstract representations.  
• Edges (Weights) – Parameters that scale signals between nodes; optimized during training.  
• Output Layer – Produces predictions (e.g., class probabilities, regression value).  
This layered design, combined with backpropagation, makes neural networks powerful universal function approximators.

**Summary**

• Deep Learning leverages multi-layer neural networks to learn complex patterns directly from raw data.  
• The history of neural networks spans perceptrons, AI winters, and revival through backpropagation.  
• The perceptron introduced learnable weights but was limited to linear problems.  
• Backpropagation enabled ANN, CNN, and RNN architectures to train effectively at scale.  
• A neural network consists of input, hidden, and output layers connected by weighted edges, forming the basis of all deep models.