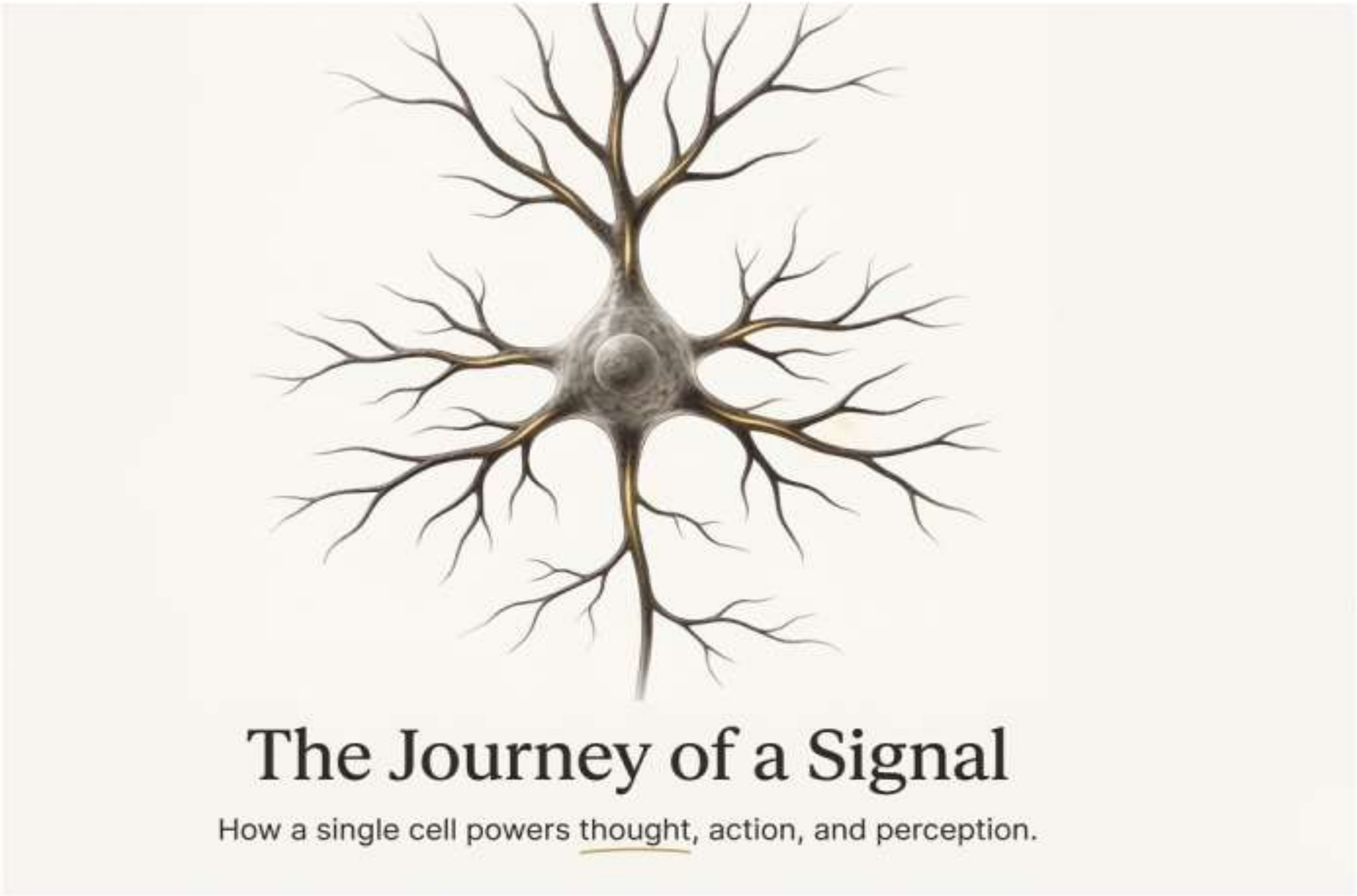
Introduction
to
Artificial Neural Network
(ANN)

[Single Layer and Multi-Layer]

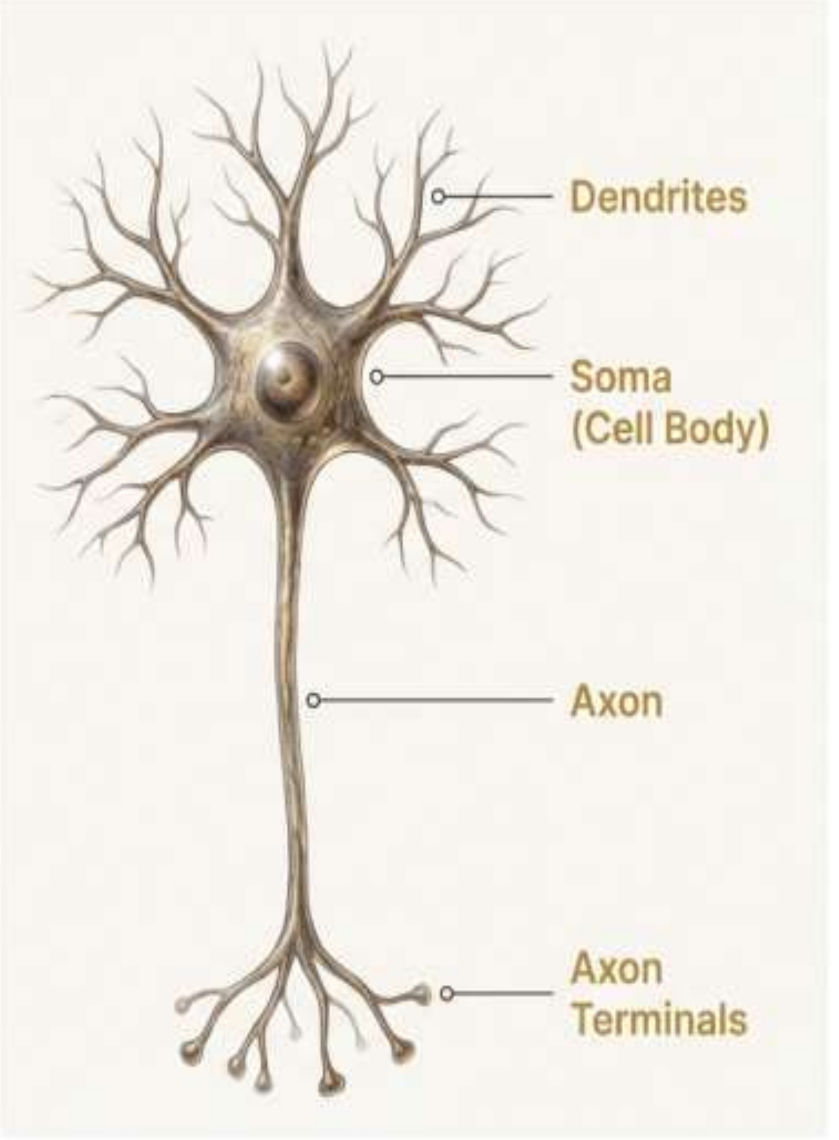
By Dr. Gopal Chandra Jana, Assistant Professor, DCSE, SSCSE

Neural Network

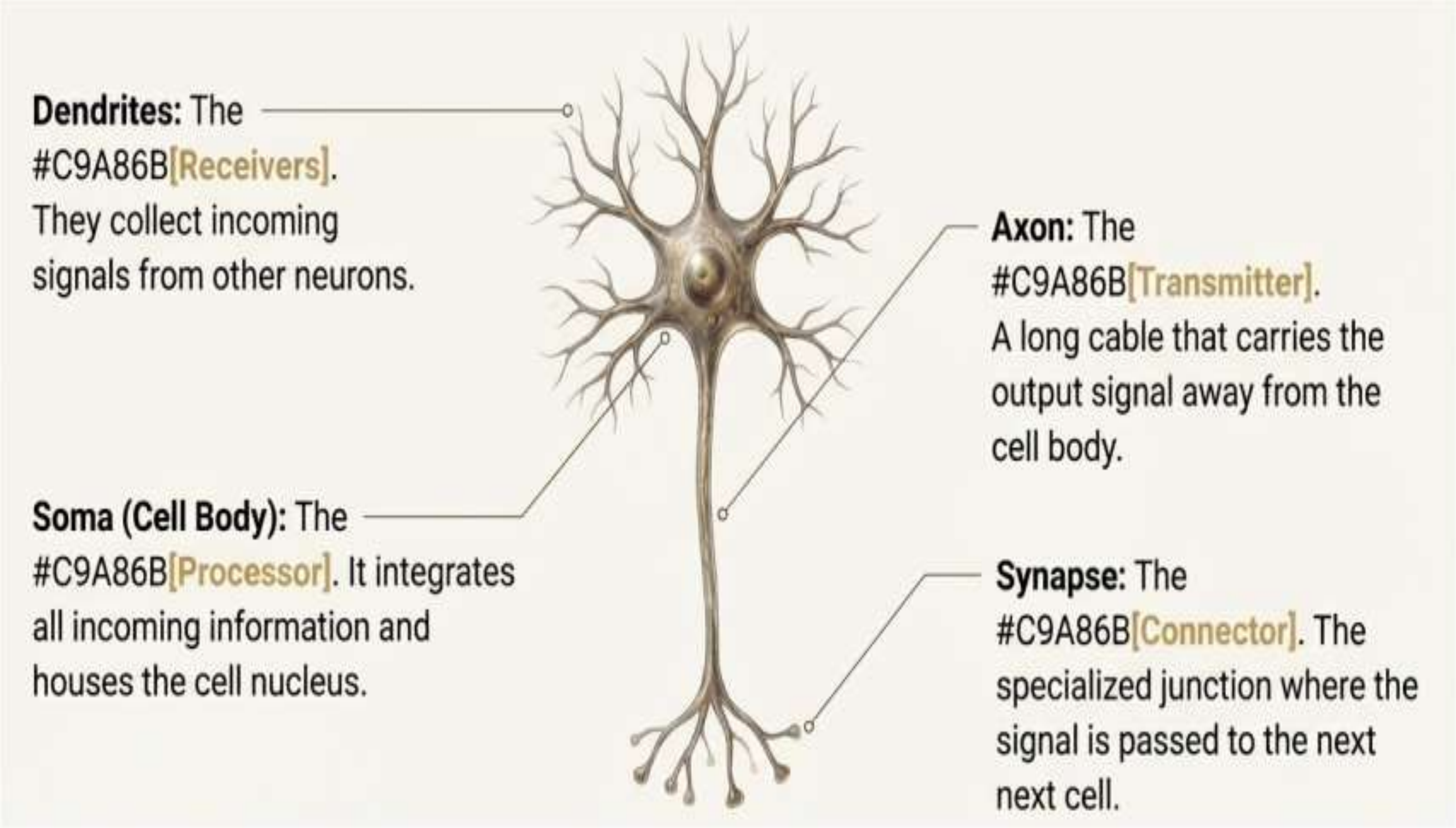


The Nervous System is Built from Billions of Specialized Cells.

- The human brain is a network of extraordinary complexity.
- Its fundamental building block is a single, elegant cell: the neuron.
- These cells are responsible for receiving, processing, and transmitting the information that underpins every aspect of our experience.
- Understanding the neuron is the first step to understanding ourselves.

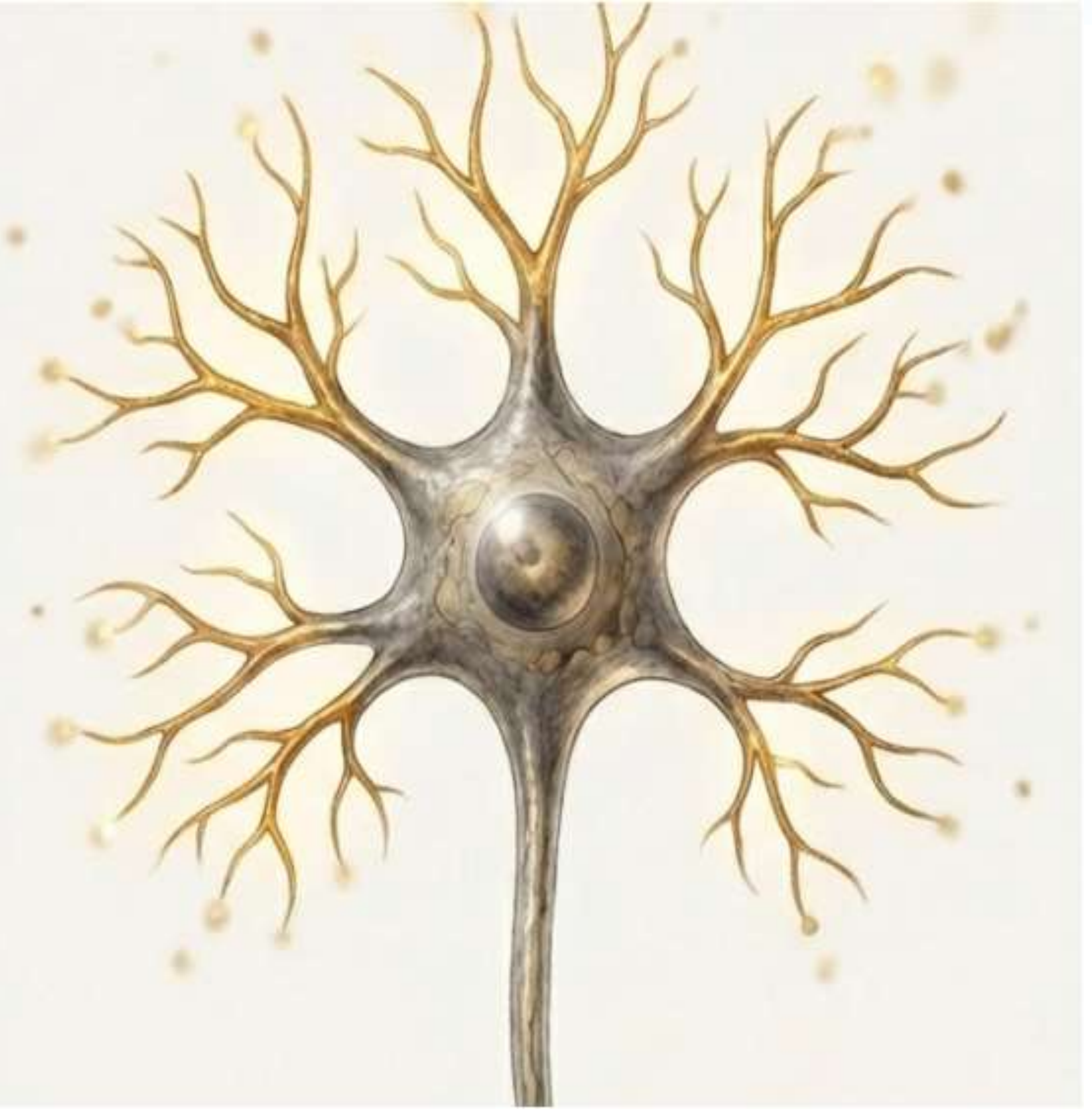


Each Part of the Neuron Plays a Crucial Role.



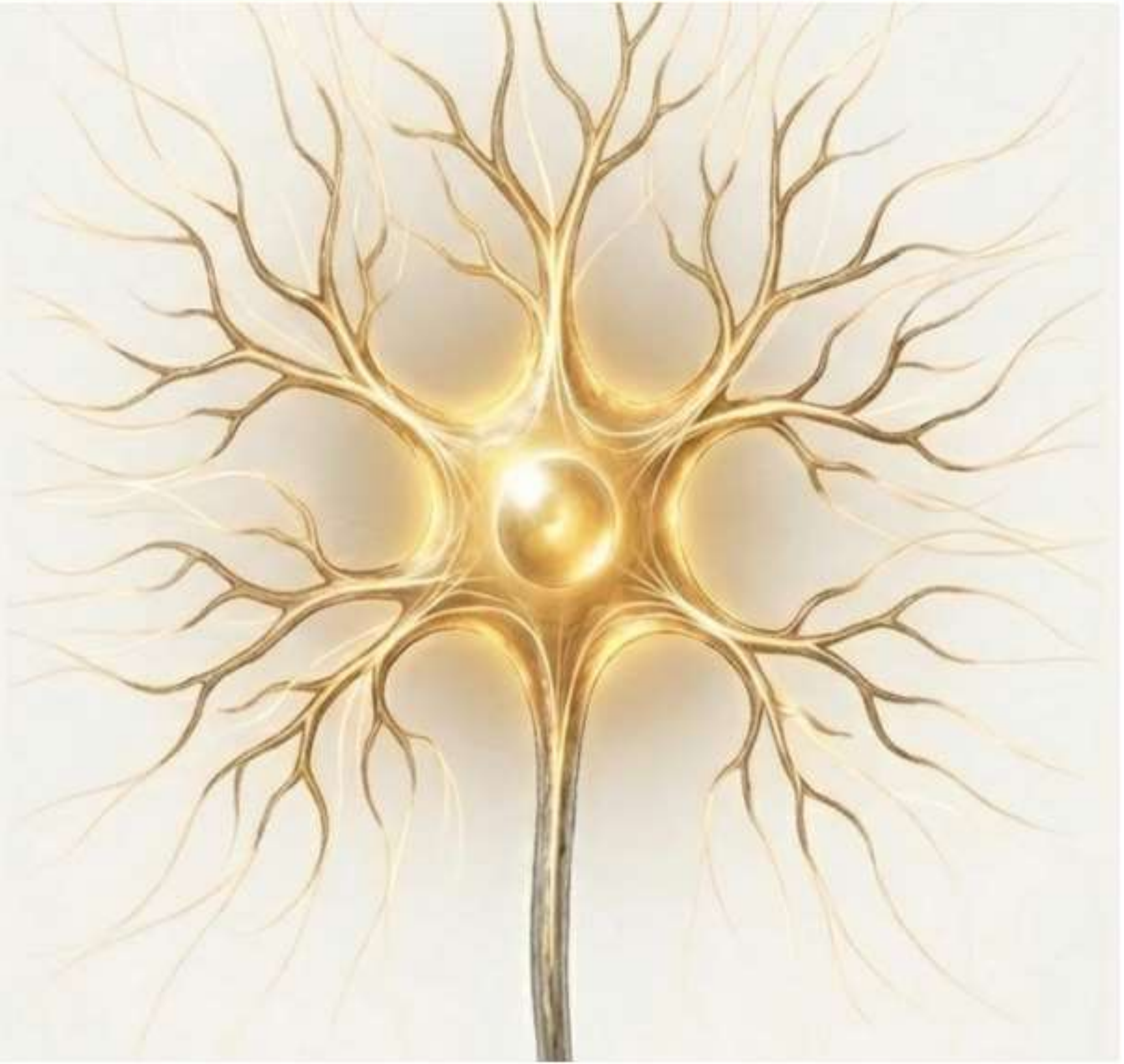
Information Arrives at the Neuron's Dendritic "Antennas".

A neuron's journey begins when it "listens" to its neighbors. Dendrites, with their vast, branching structures, act as sensitive antennas. They receive chemical signals (neurotransmitters) from thousands of other neurons, converting them into tiny electrical currents.



The Cell Body Integrates a Chorus of Incoming Messages.

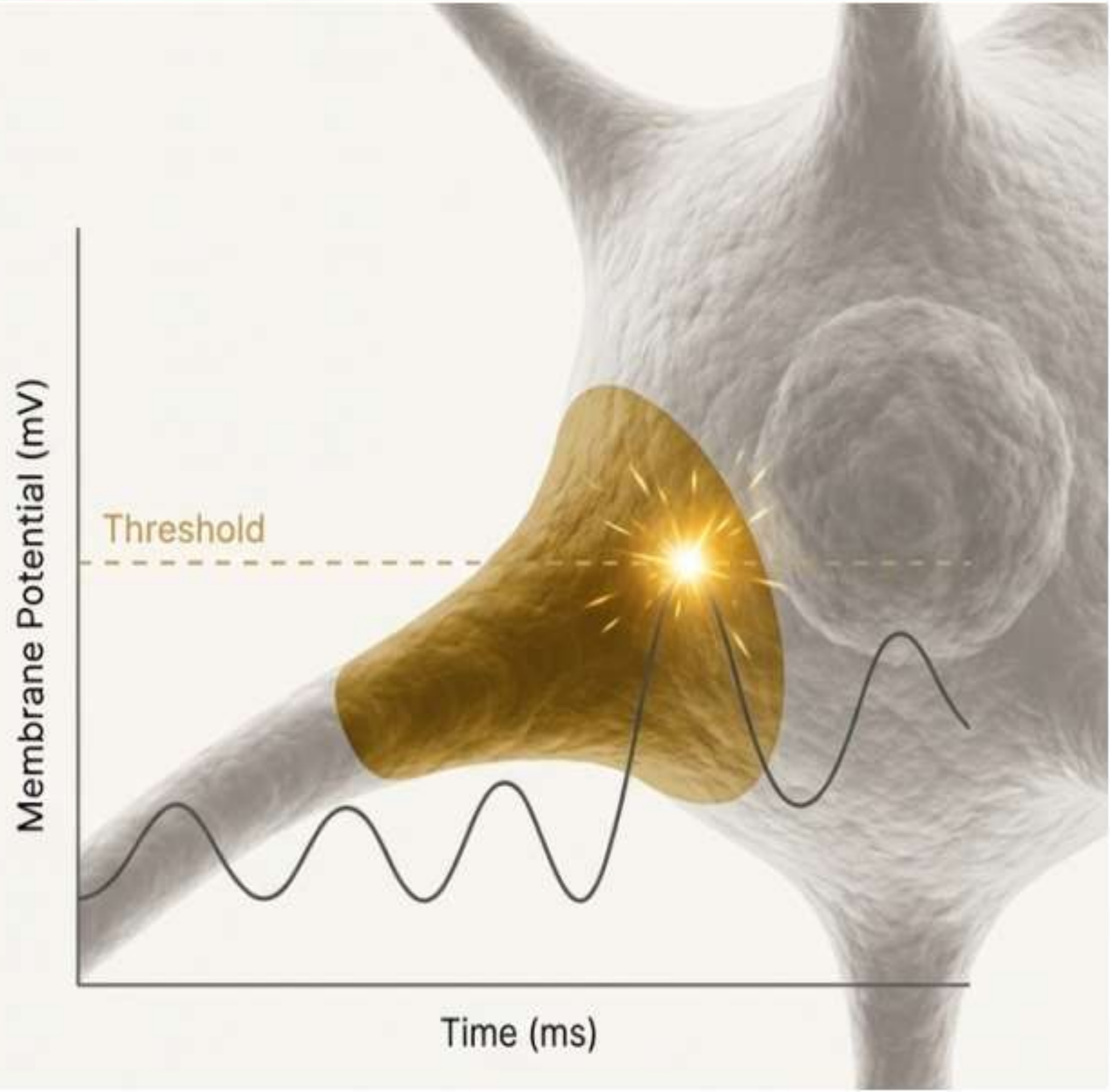
The small electrical currents from the dendrites flow to the cell body, or soma. Here, all these signals—some excitatory (pushing the neuron to fire) and some inhibitory (holding it back)—are summed together. The soma also contains the cell nucleus, the vital center that maintains the neuron's life and function.



A Critical Threshold Must Be Reached for the Neuron to Fire

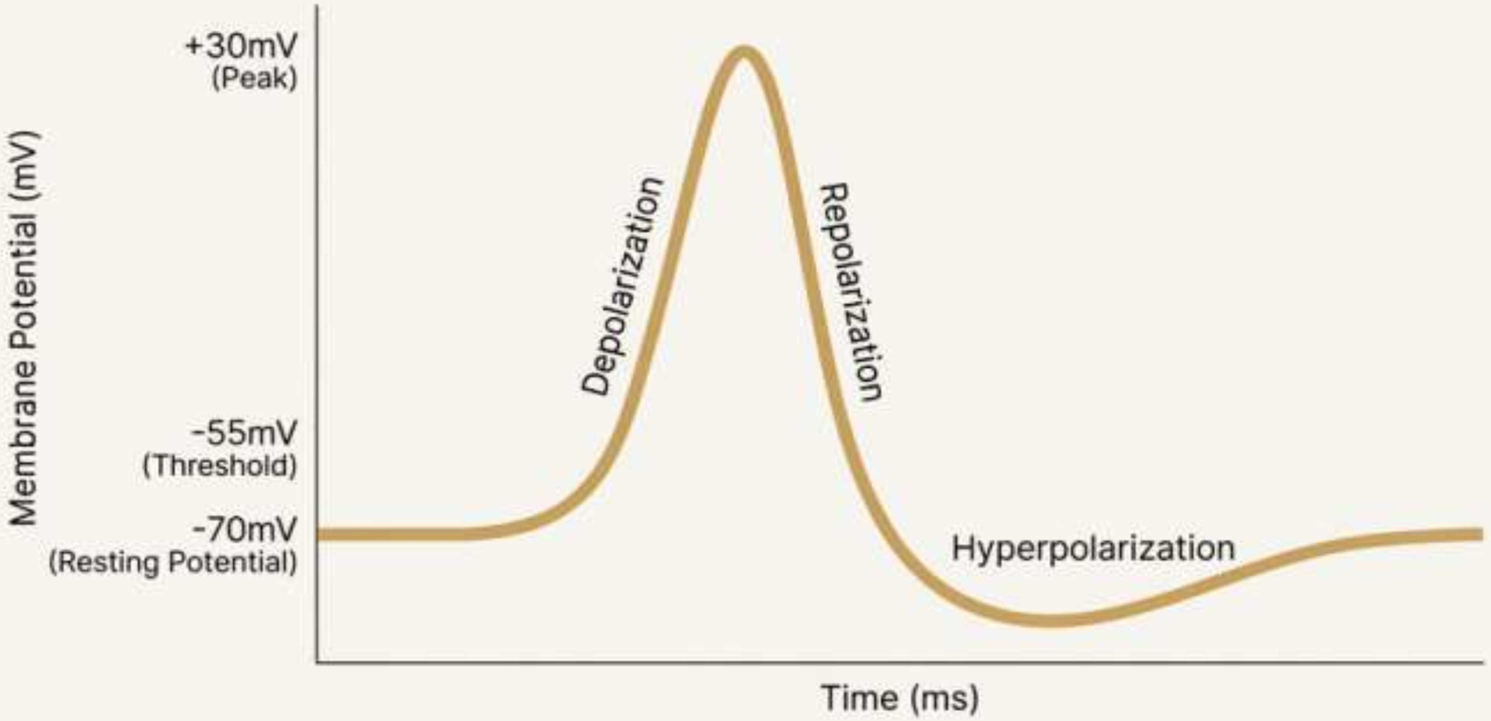
At the junction between the soma and the axon lies a specialized region called the **axon hillock**. This area acts as a decision-maker.
If the total integrated signal in the soma reaches a specific voltage threshold, the **axon hillock** initiates an unstoppable chain reaction.

It is the neuron's trigger.



The Journey

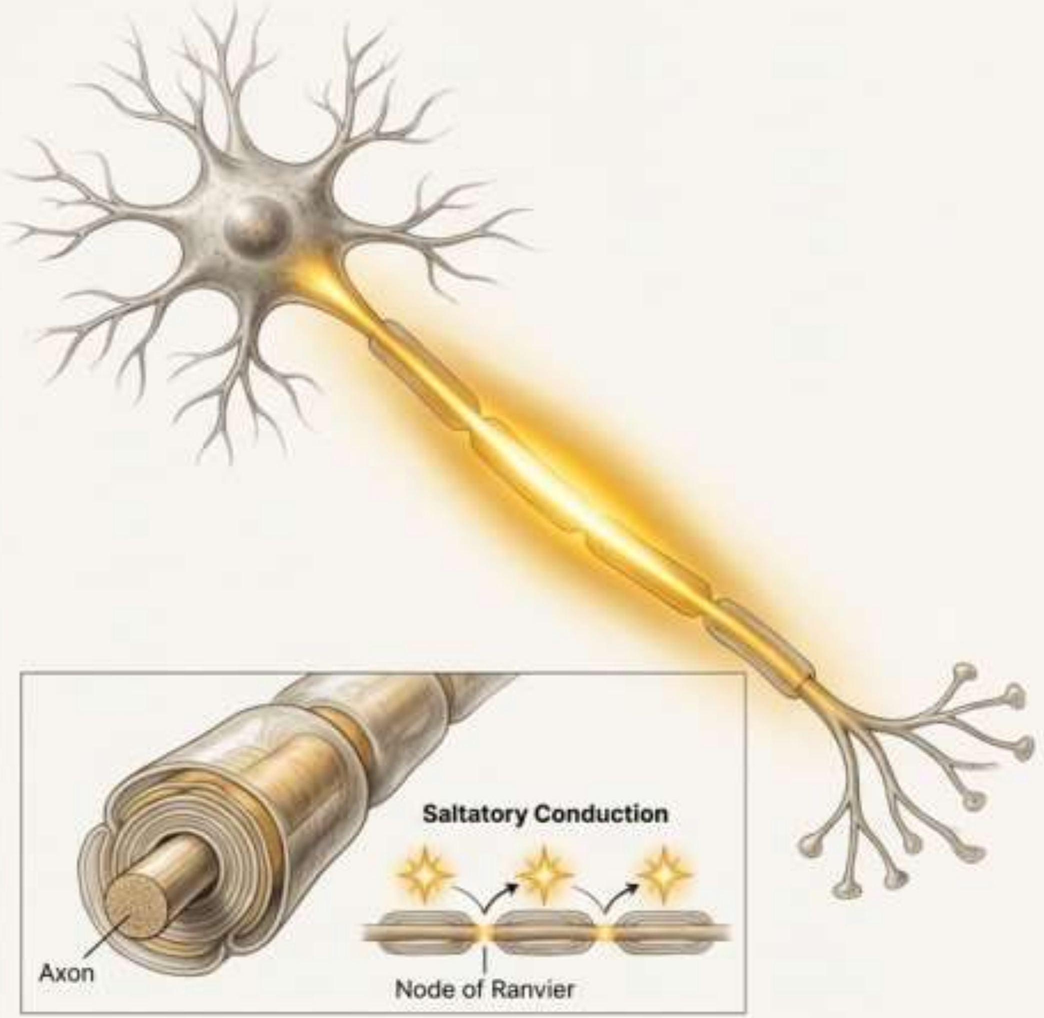
The Action Potential: An "All-or-None" Electrical Impulse.



Once the threshold is crossed, the neuron fires an action potential. This is not a graded signal; it is a massive, rapid, all-or-none electrical impulse of a fixed size and duration. The neuron doesn't fire 'harder' or 'softer'—it either fires at full strength, or it does not fire at all.

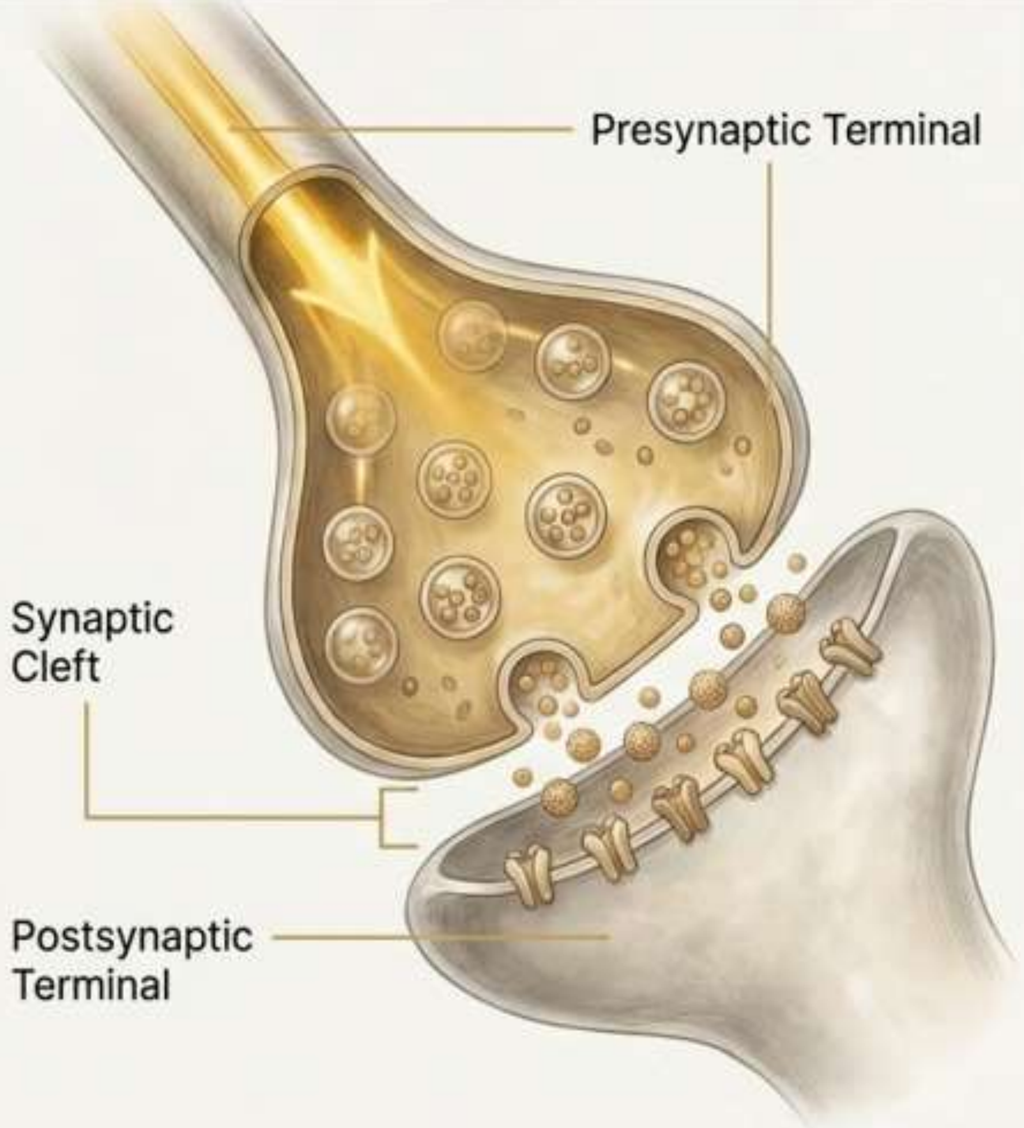
The Signal Races Down the Axon, an Insulated Cellular Cable

The action potential propagates itself down the length of the axon. In many neurons, the axon is wrapped in a fatty **myelin sheath**, which acts like insulation on a wire. This insulation forces the signal to "jump" between gaps, dramatically increasing the speed and fidelity of transmission.



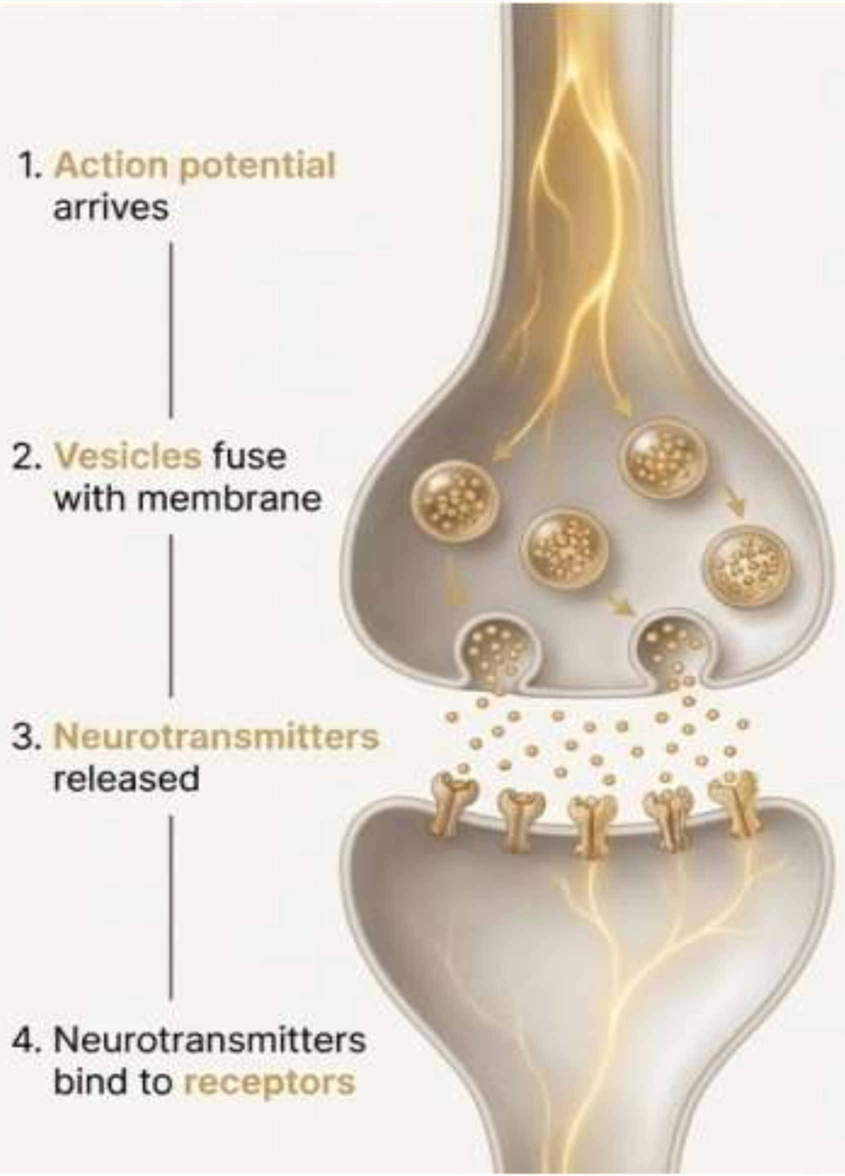
The Signal Reaches the Synapse, the Bridge to the Next Neuron.

The axon terminates in a structure called the synapse. This is not a physical connection but an incredibly narrow gap—the synaptic cleft—that separates one neuron from the next. To cross this bridge, the signal must transform from an electrical one into a chemical one.

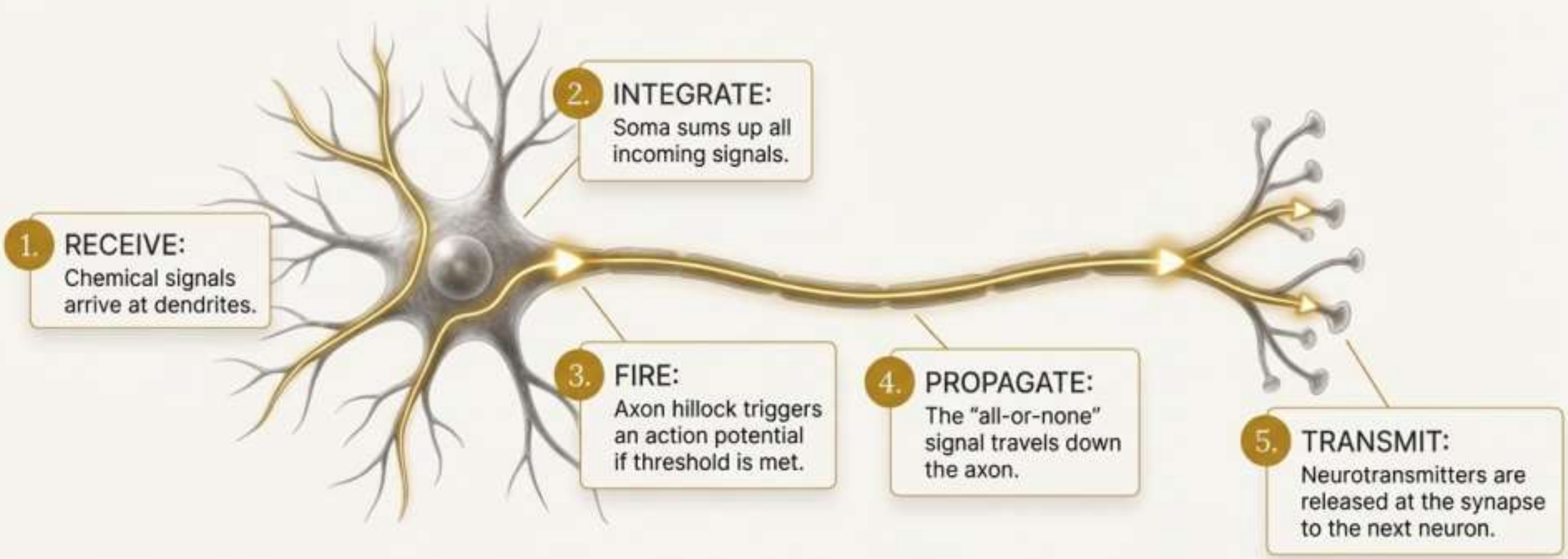


Communication Across the Synapse is a Chemical Conversation.

When the action potential arrives, it triggers the release of **chemical messengers** called neurotransmitters from the axon terminal. These molecules diffuse across the tiny synaptic cleft and bind to receptors on the next neuron's dendrite, initiating a new electrical current and beginning the journey all over again.



From Input to Output: A Complete Neural Signal Pathway



This Simple Journey, Multiplied by Bilions, Creates a Biological Computer.

The process of receiving, integrating, and transmitting signals is the fundamental algorithm of the nervous system. While a single neuron's task is simple, the coordinated firing of billions of neurons in a complex, interconnected network gives rise to every thought, memory, and action.

Our Understanding Was Forged Over a Century of Discovery.

The neuron was not always understood as an individual unit. Early theories proposed a continuous "nerve net." It was the pioneering work of scientists like **Santiago Ramón y Cajal**, using staining techniques developed by Camillo Golgi, that established the "Neuron Doctrine"—the principle that the nervous system is made up of discrete, individual cells.

c. 1873
Golgi's Stain reveals the full structure of a neuron for the first time.

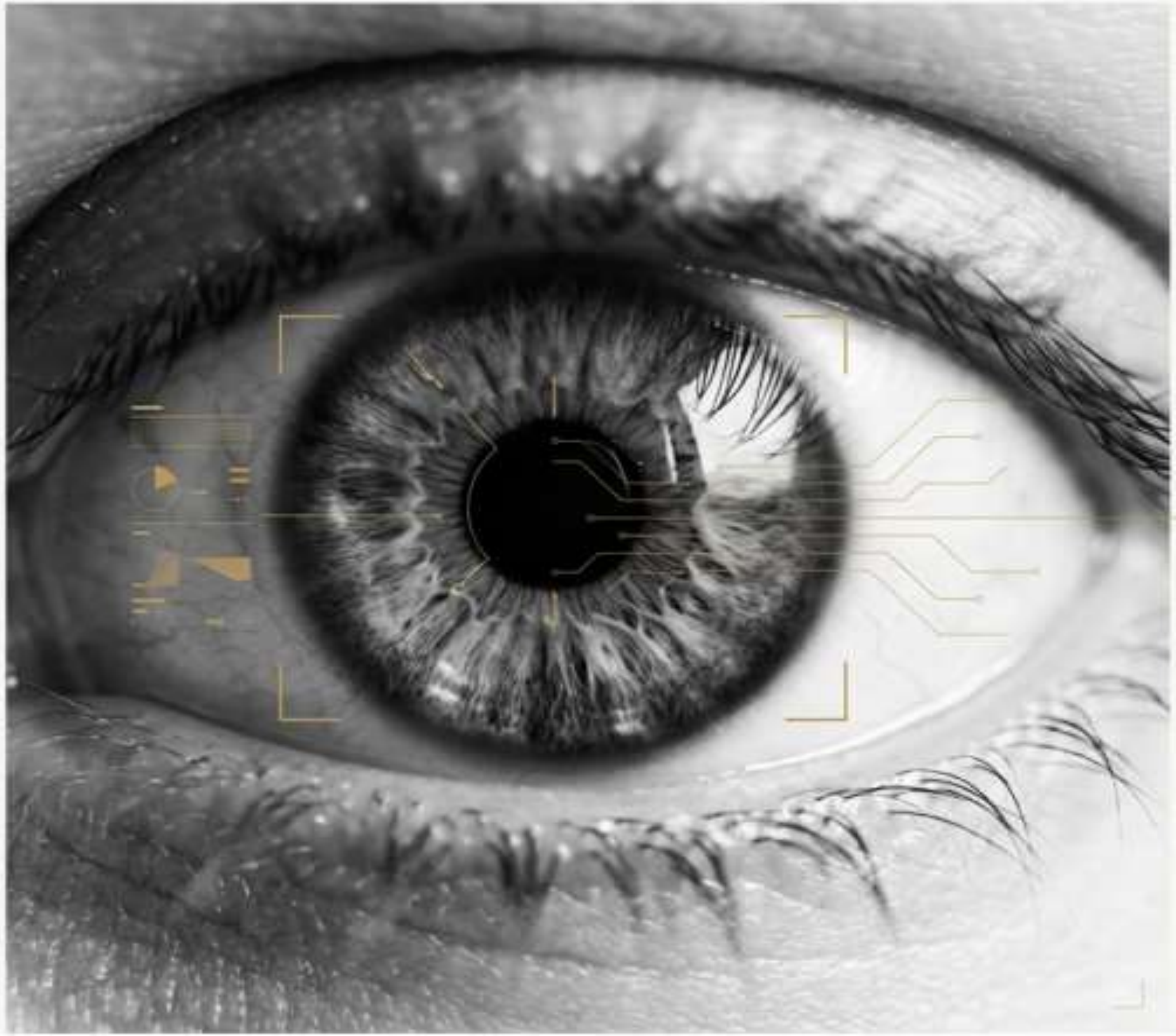
c. 1891
Ramón y Cajal proposes the Neuron Doctrine, correctly identifying neurons as individual cells.

c. 1952
Hodgkin & Huxley model the action potential, explaining the electrical journey.

c. 1873 — c. 1891 — c. 1952

The Neuron is a Masterpiece of Biological Engineering.

The journey of a single signal reveals a cell of breathtaking efficiency and elegance. From its ability to integrate vast streams of information to its high-speed, high-fidelity transmission, the neuron is the physical foundation of consciousness—a perfect fusion of biology and computation.



The Biological Supercomputer

The human brain is a highly complex, nonlinear, and parallel information-processing system that performs complex tasks orders of magnitude faster than the most powerful digital computers.

100–200 milliseconds

The time it takes for the brain to accomplish a complex perceptual task, like recognizing a familiar face in an unfamiliar scene.

SUPPORTING EXAMPLE

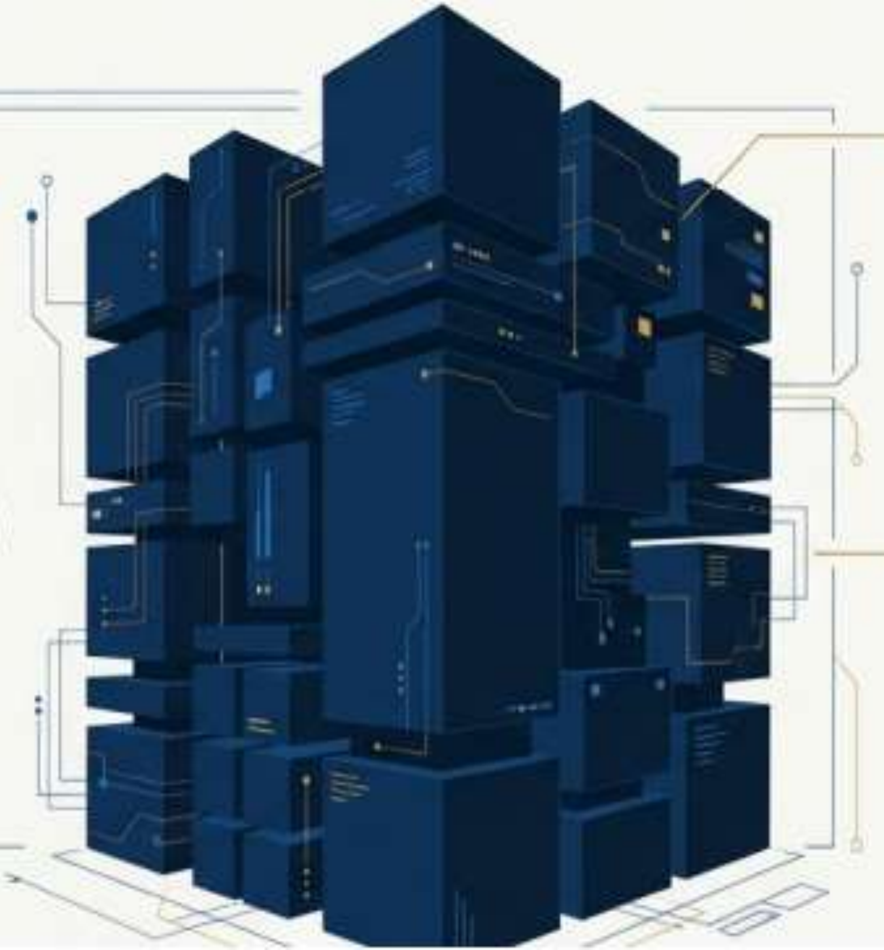
The sonar of a bat provides another example. Its brain, the size of a plum, processes echoes to determine a target's distance, velocity, size, and location with a success rate that would be "the envy of a radar or sonar engineer."

Unmatched Computational Efficiency

Human Brain

10^{-16} Joules

Energy per operation per second. The value for the best computers is orders of magnitude larger.



Scale

Compensates for slower neuron switching speeds (milliseconds vs. nanoseconds) with massive parallelism.

Numbers

- 10 Billion Neurons (estimated)
- 60 Trillion Synapses (estimated)

The Secret is Plasticity

"Plasticity permits the developing nervous system to adapt to its surrounding environment." In an adult brain, this is achieved through:

1. The creation of new synaptic connections.
2. The modification of existing synapses.

Experience is built up over time. The brain isn't static; it continuously 'hardwires' itself based on interaction with the world. This adaptability is the key feature we seek to emulate in artificial neural networks.

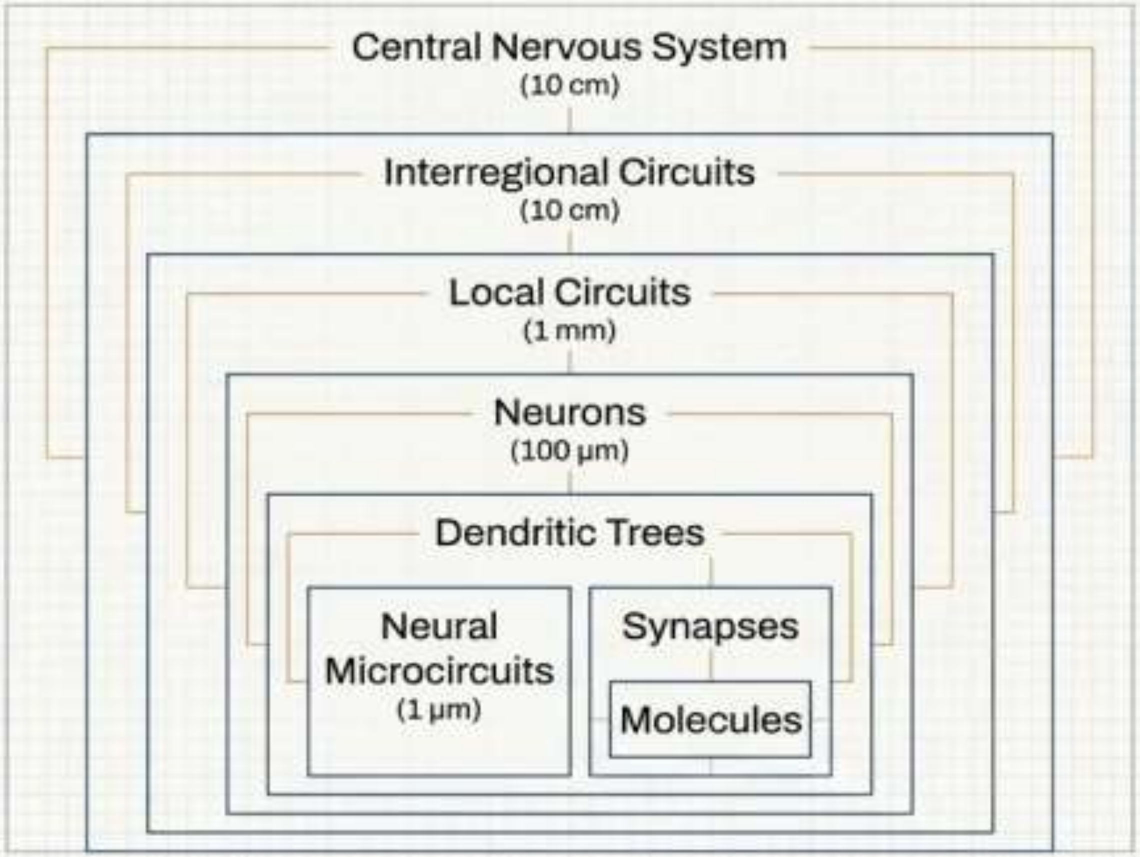
Defining the Artificial Counterpart

> *A neural network is a massively parallel distributed processor made up of simple processing units that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:*

1. **Learning Process:** Knowledge is acquired by the network from its environment.
2. **Knowledge Storage:** Interneuron connection strengths, known as *synaptic weights*, are used to store the acquired knowledge.

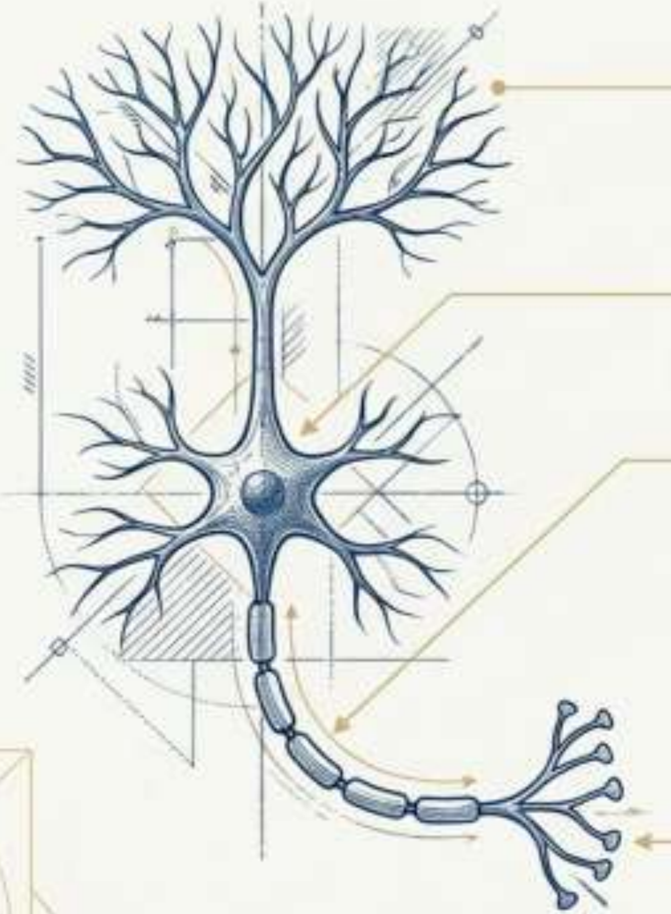
The procedure used to modify these weights is called a **learning algorithm.**

Levels of Organization in the Brain



This deep structural hierarchy is a unique characteristic of biological computation. Our artificial models are, by comparison, extremely primitive but follow a similar layered philosophy.

The Fundamental Unit: The Neuron



Dendrites

The "receptive zones," resembling a tree, that receive inputs from other neurons (can be 10,000+ contacts).

Cell Body

Processes the incoming signals from the dendrites.

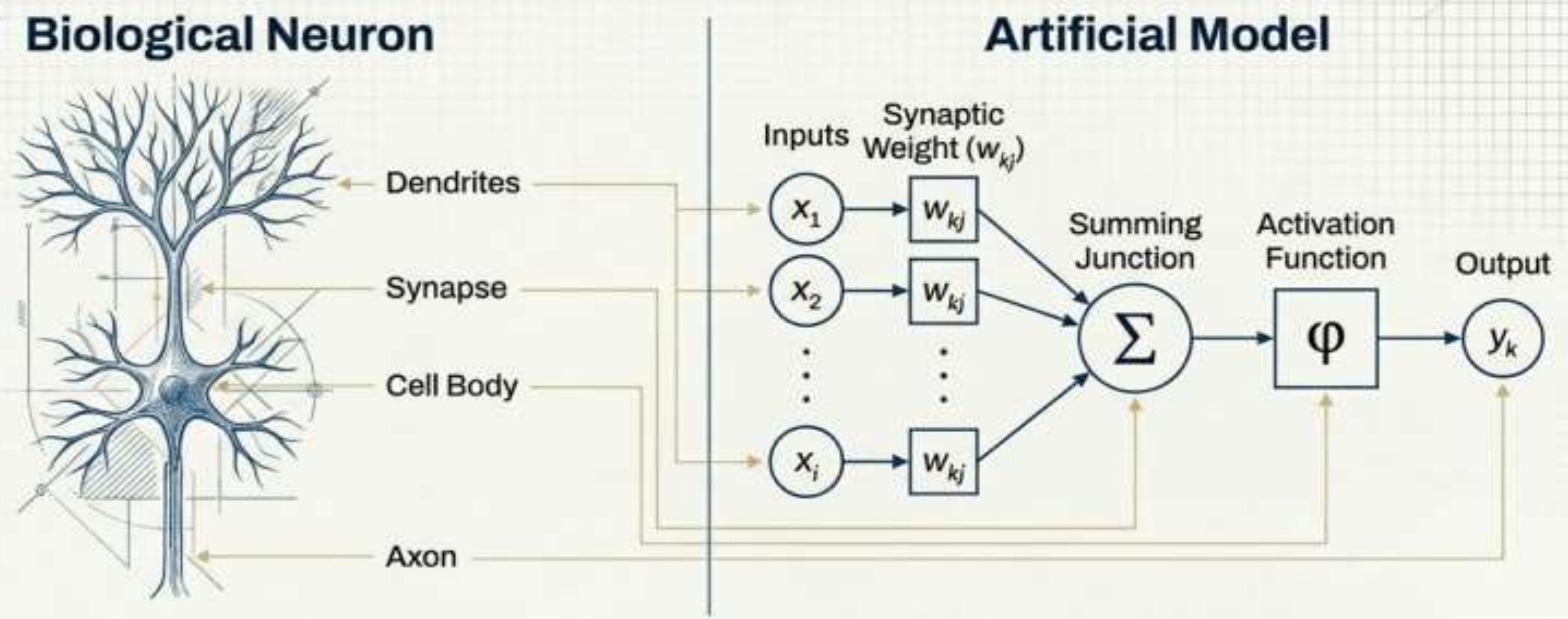
Axon

The "transmission line" that propagates the output signal (as an *action potential* or *spike*) to other neurons.

Synapse

The junction where signals are transmitted (usually chemically) from one neuron to another. This is where *excitation* or *inhibition* occurs.

From Biological Blueprint to Artificial Model



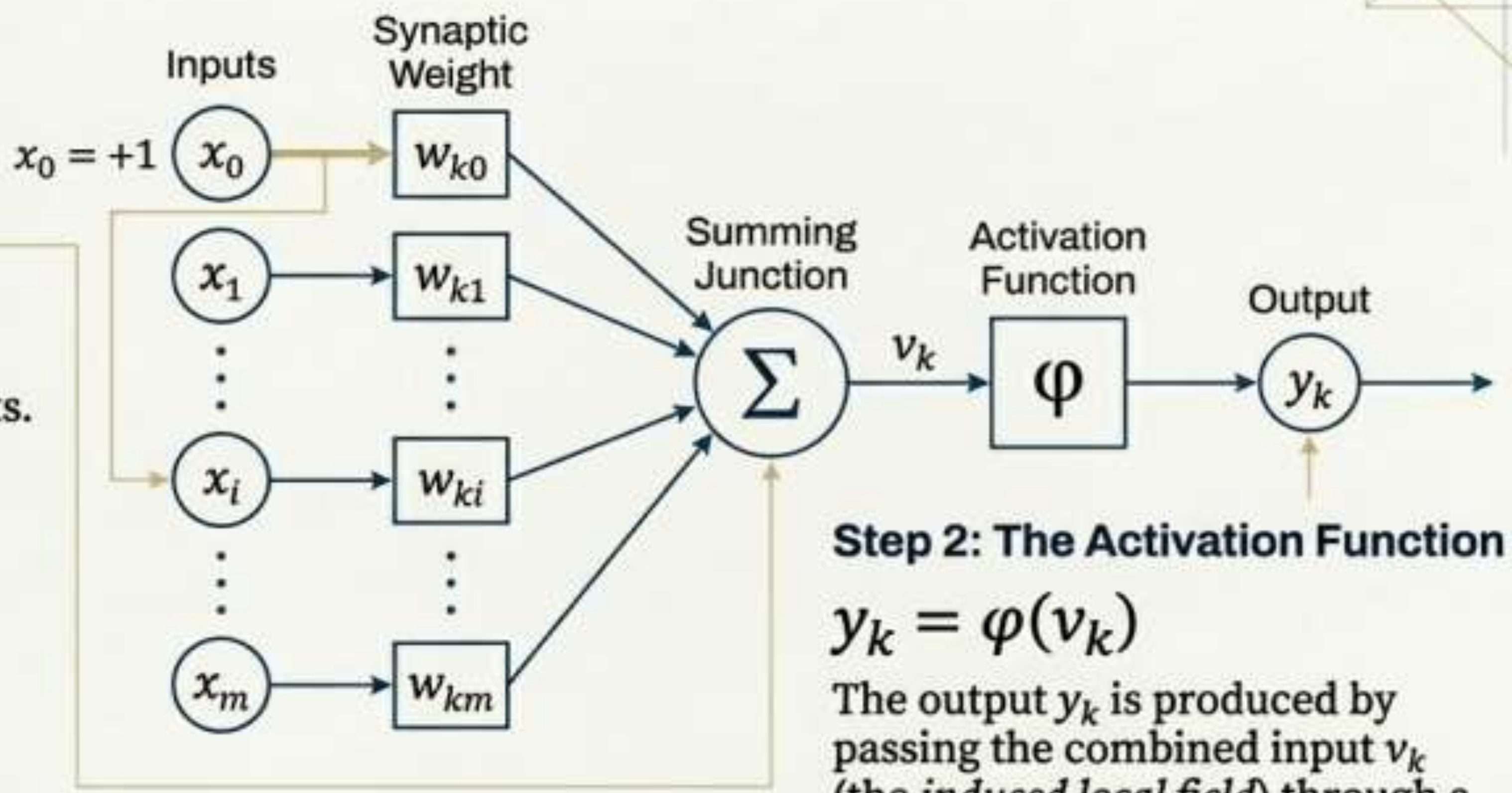
The artificial neuron is an information-processing unit that is fundamental to the operation of a neural network. It is a simplified mathematical model of its biological counterpart.

The Mathematics of a Neuron

Step 1:
The Linear Combiner

$$v_k = \sum_{j=0}^{m} w_{kj} x_j$$

The neuron computes a weighted sum of its m inputs. The *bias* (b_k or w_{k0}) acts as an external parameter that increases or lowers the net input, allowing the activation function's response to be shifted.



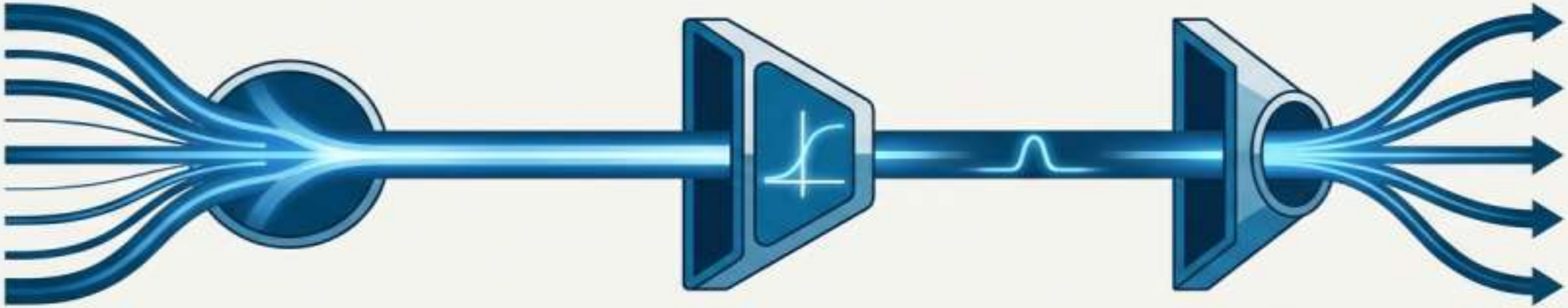
Step 2: The Activation Function

$$y_k = \varphi(v_k)$$

The output y_k is produced by passing the combined input v_k (the *induced local field*) through a nonlinear *activation function* φ.

The Three Functions of an Artificial Neuron

An artificial neuron processes data through a sequence of three distinct mathematical functions.



1. Propagation Function

Aggregates inputs. The most popular form is the weighted sum.

$$net_j = \sum (o_i * w_{i,j})$$

Analogy: Weights are like volume dials for each input's importance.

2. Activation Function

Introduces non-linearity. It determines the neuron's "switching status" based on the network input and a threshold.

$$a_j(t) = fact(net_j(t), a_j(t-1), \Theta_j)$$

3. Output Function

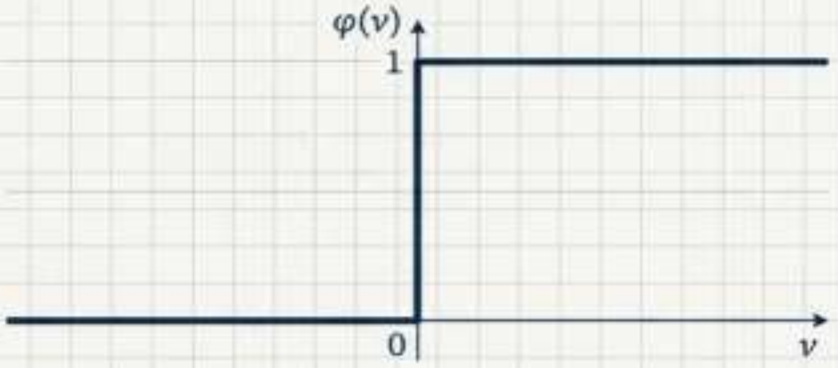
Transmits the result. It calculates the final value sent to other neurons. This is often the identity function.

$$O_j = a_j$$

The Activation Function: Deciding the Output

The activation function (or "squashing function") limits the amplitude of the neuron's output to a finite range (e.g., [0, 1] or [-1, 1]).

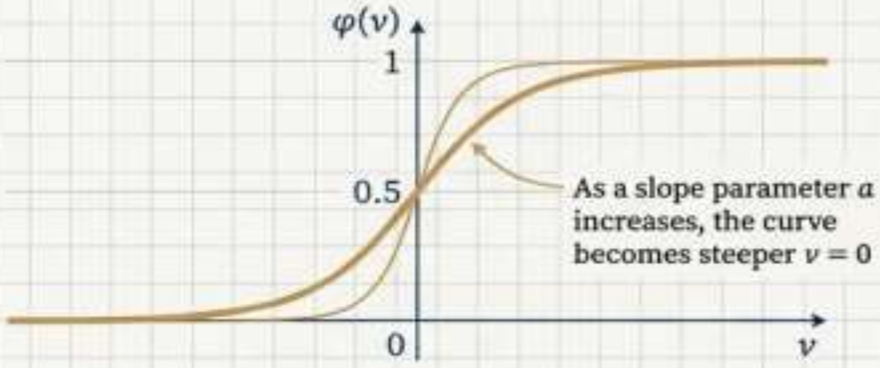
1. Threshold Function (Heaviside function)



An "all-or-none" function. The output is 1 if the input v v is non-negative, and 0 otherwise.

This forms the basis of the **McCulloch-Pitts model** of a neuron.

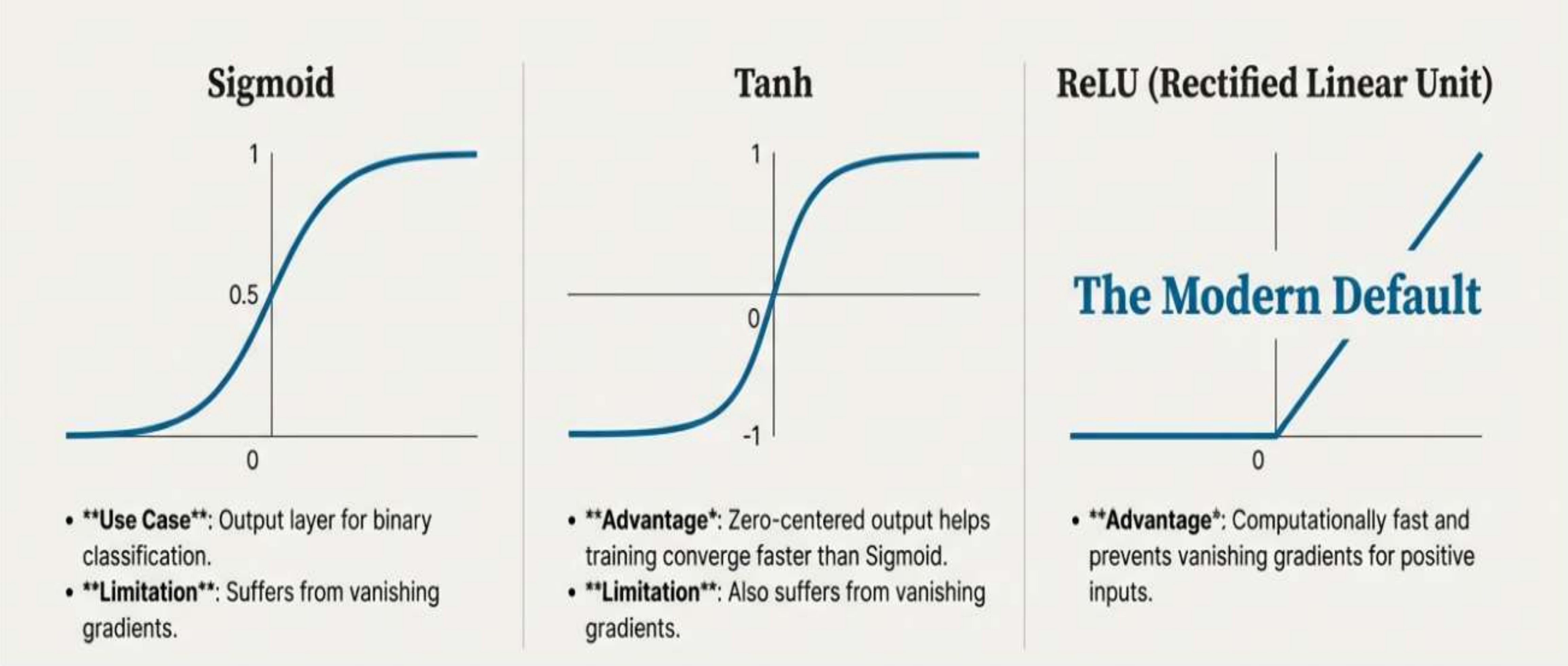
2. Sigmoid Function



An "S-shaped," strictly increasing function that is continuous and differentiable. The logistic function $\varphi(v) = 1 / (1 + \exp(-av))$ is a common example.

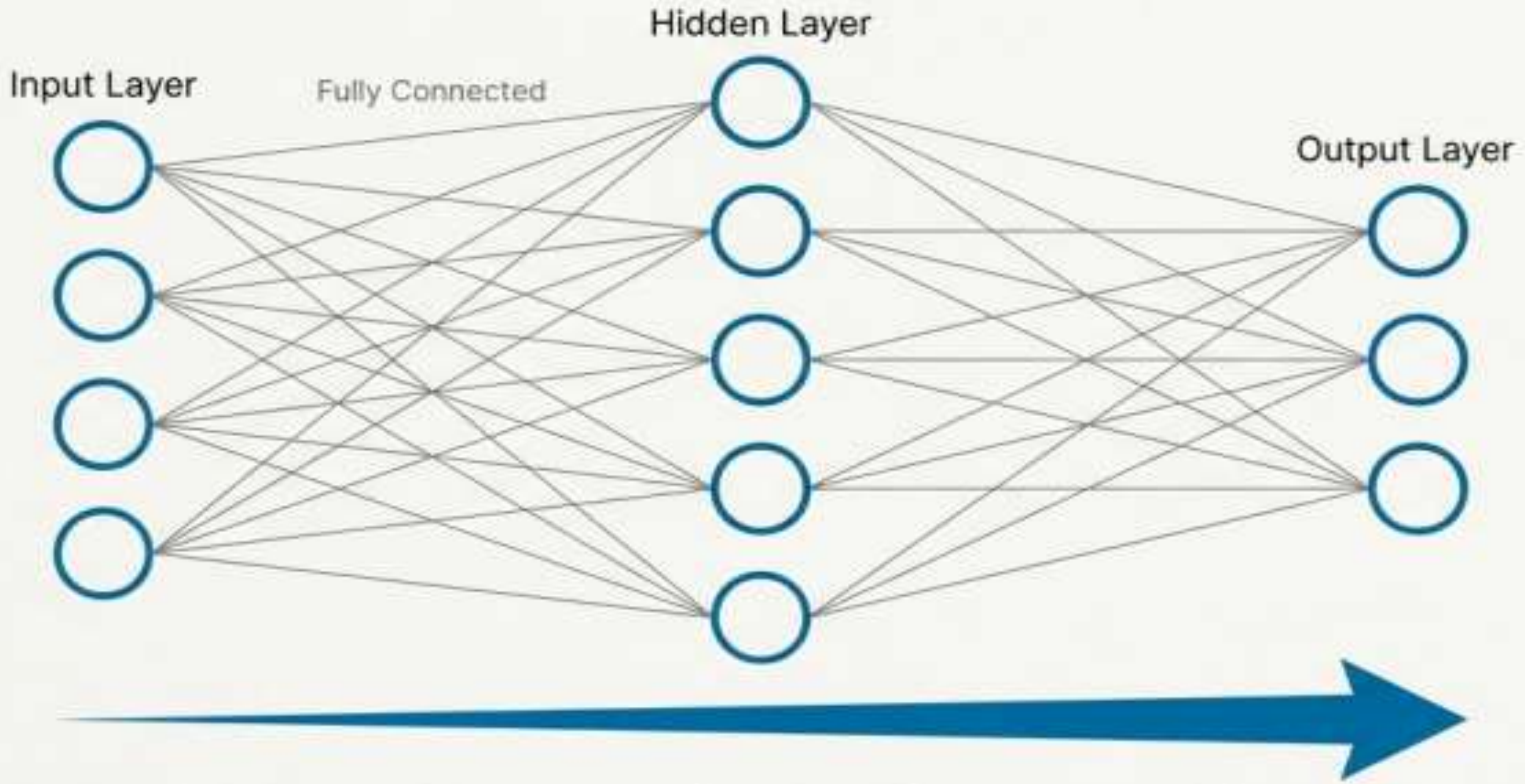
The parameter a controls the slope. As a approaches infinity, the sigmoid function becomes a threshold function.

- Without activation functions, a multi-layer neural network would be no more powerful than a simple linear model. These functions allow the network to learn complex, non-linear patterns in data.



Assembling the Network: From Neurons to Layers.

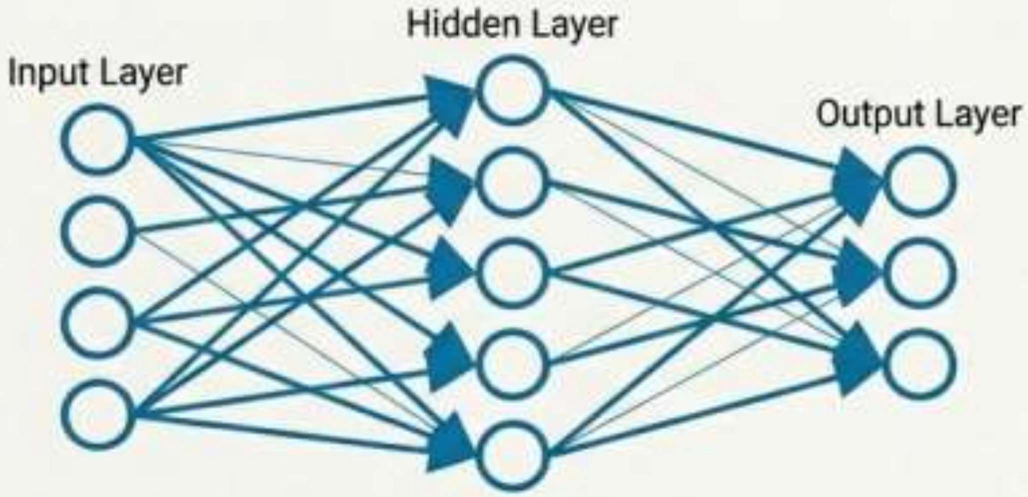
The true power of neural networks emerges when neurons are organized into layers. Data flows sequentially from the input layer, through one or more hidden layers, to the output layer.



Feedforward Neural Network: A network where connections do not form cycles and information moves in one direction.

Two Architectural Blueprints: FNN vs. RNN.

Feedforward Neural Networks (FNN)



Key Idea: No memory of previous inputs. Connections do not form cycles.

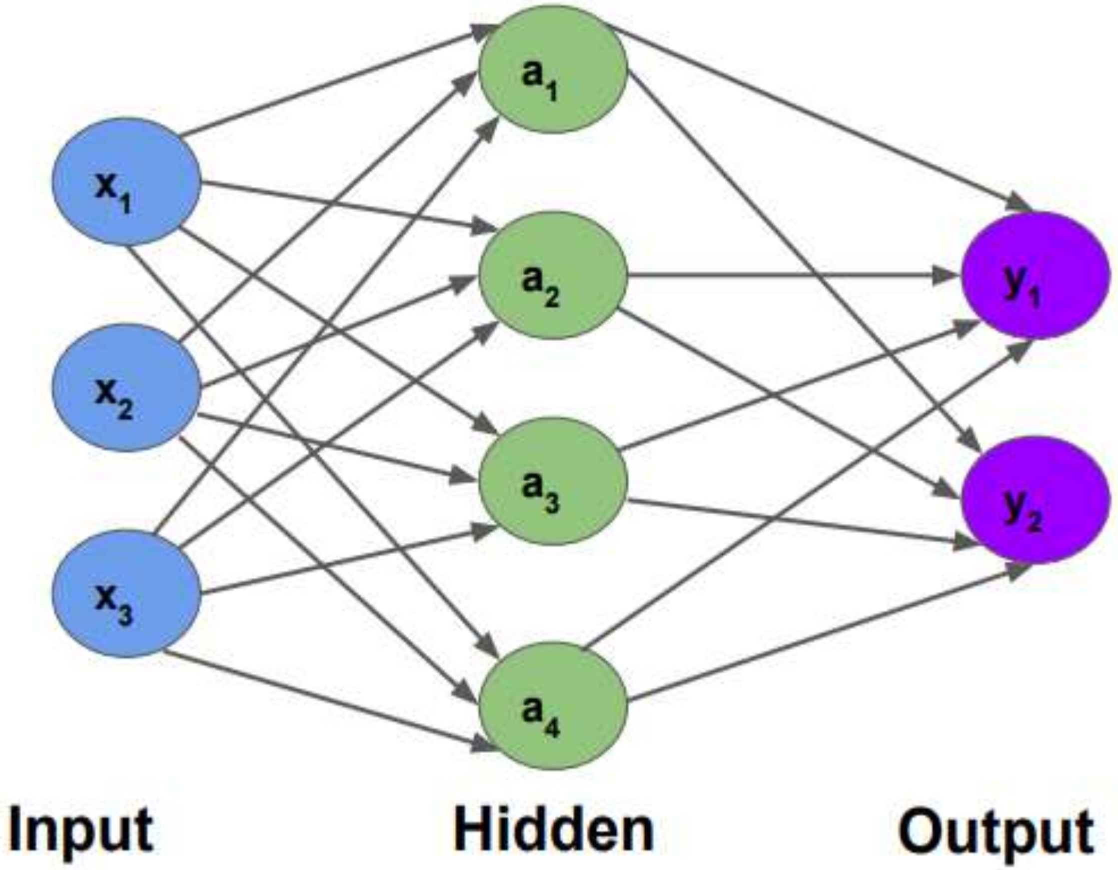
Best For: Static data where inputs are independent (e.g., image classification, credit scoring).

Recurrent Neural Networks (RNN)

Key Idea: Has memory via hidden states. Ideal for sequential data where context matters.

Best For: Sequential data where order is critical (e.g., language translation, stock price prediction).

Feed-forward Neural Network



$$a_1 = f(W_{11}x_1 + W_{12}x_2 + W_{13}x_3)$$

.

.

$$a_3 = f(W_{31}x_1 + W_{32}x_2 + W_{33}x_3)$$

$$y_1 = f(U_{11}a_1 + U_{12}a_2 + U_{13}a_3 + U_{14}a_4)$$

$$f(x) = \frac{1}{1+e^{-x}}$$

Matrix form:

$$z_1 = Wx$$

$$a = f(z_1)$$

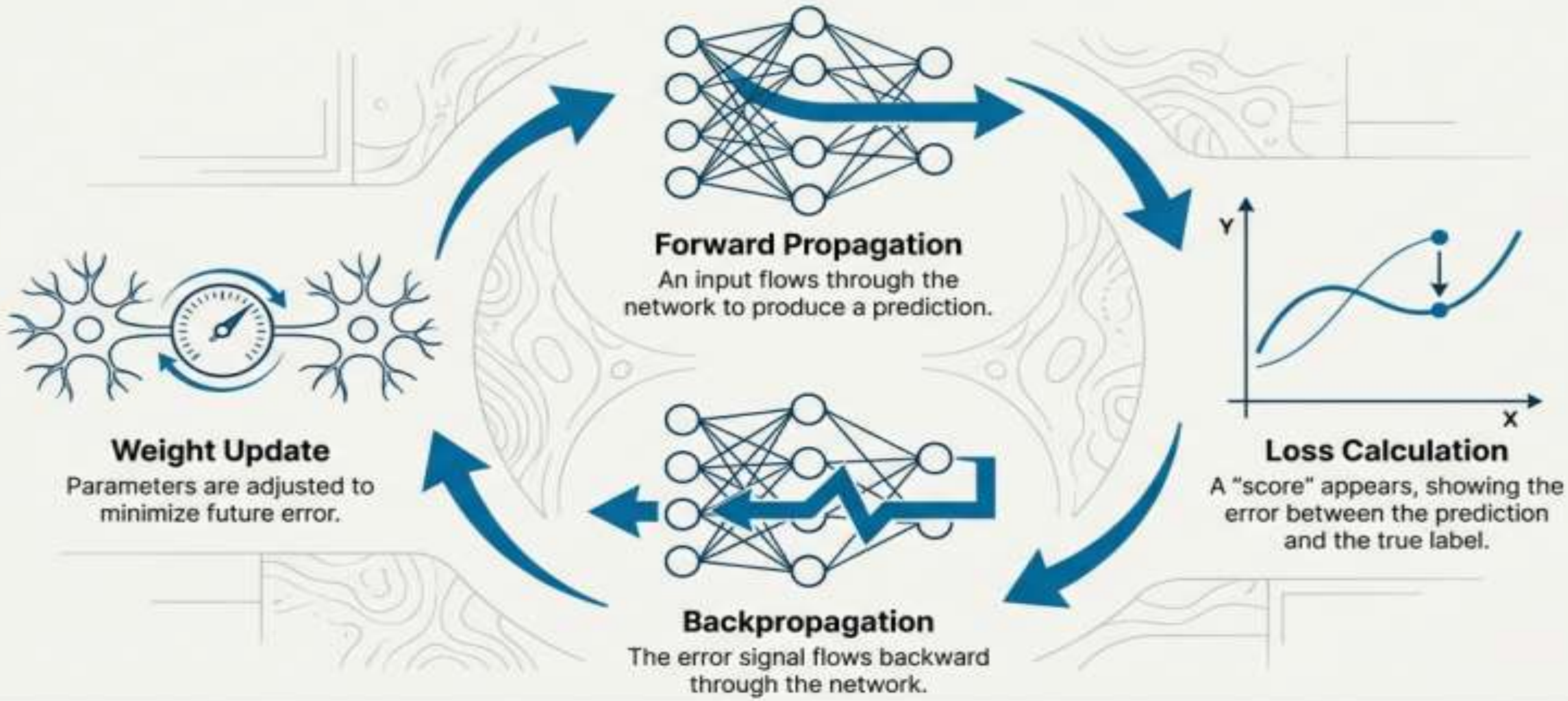
$$z_2 = Ua$$

$$y = f(z_2)$$

where $x \in \mathbb{R}^{d_i}$, $W \in \mathbb{R}^{d_1 \times d_i}$, $a \in \mathbb{R}^{d_1}$,
$U \in \mathbb{R}^{d_o \times d_1}$, $y \in \mathbb{R}^{d_o}$

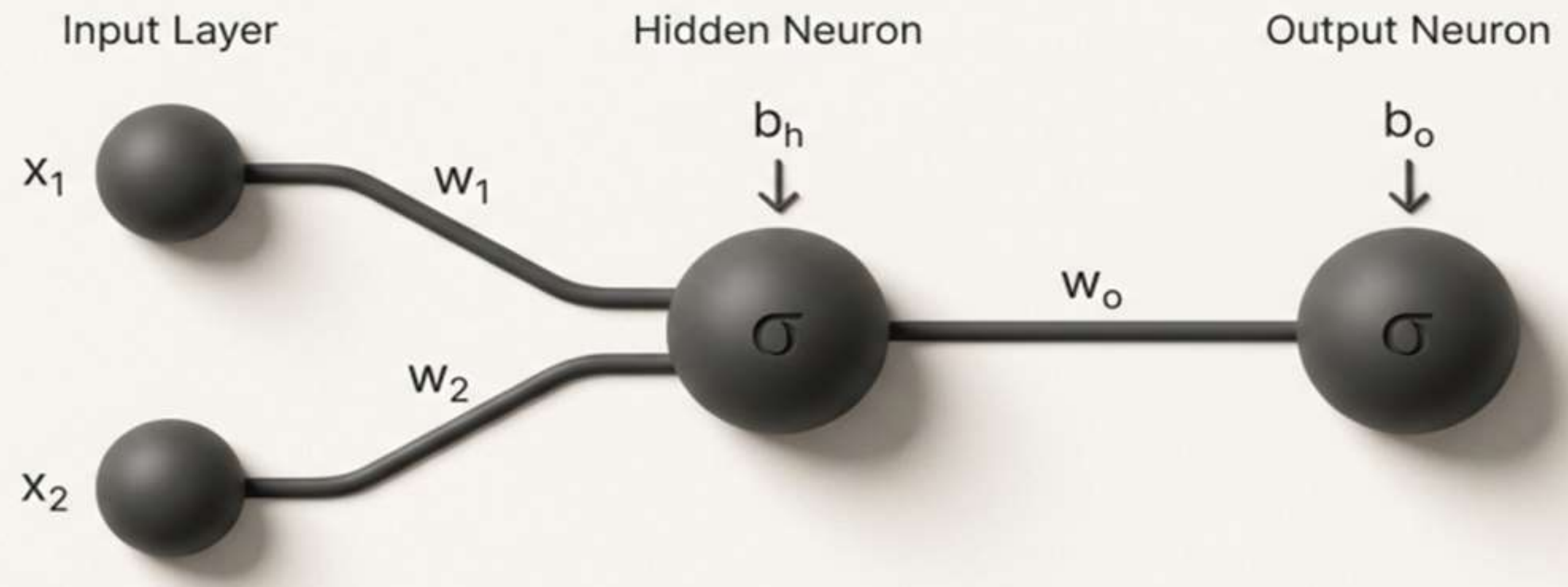
The Learning Loop

A network learns by repeating a four-step process. This cycle adjusts the network's weights and biases to progressively reduce the difference between its predictions and the true answers.



Time Solve a Numerical Problem by Considering the ANN workflow

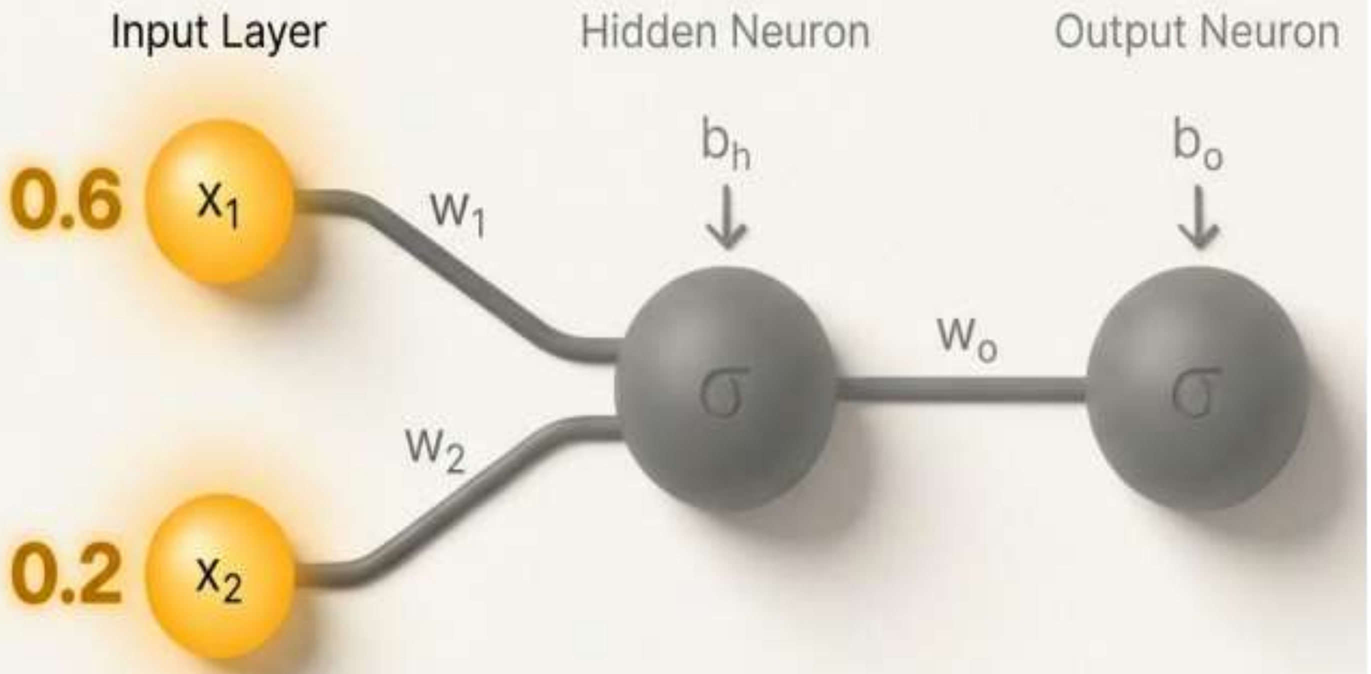
Our Map: The Network Architecture



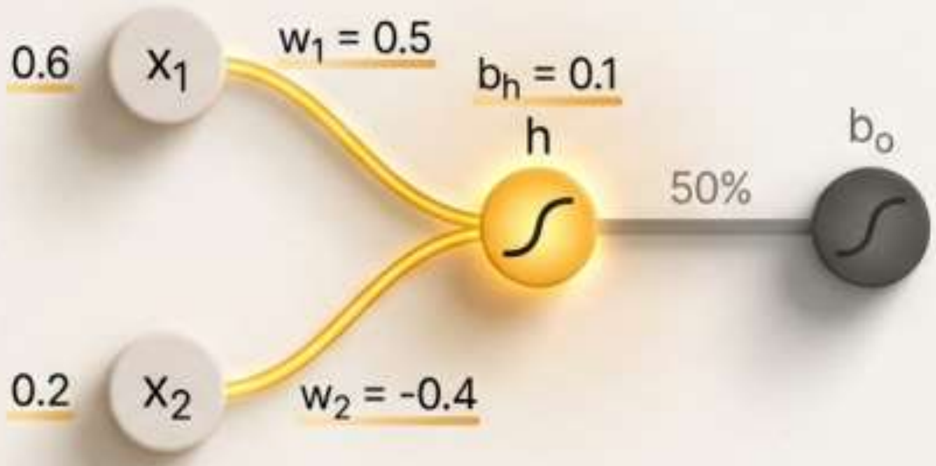
The Journey Begins with Our Inputs

Input x1 = 0.6
Input x2 = 0.2

These two signals are about to travel to the hidden neuron.



Stop 1: The Hidden Neuron Aggregates the Signals



$$z_h = (x_1 \times w_1) + (x_2 \times w_2) + b_h$$

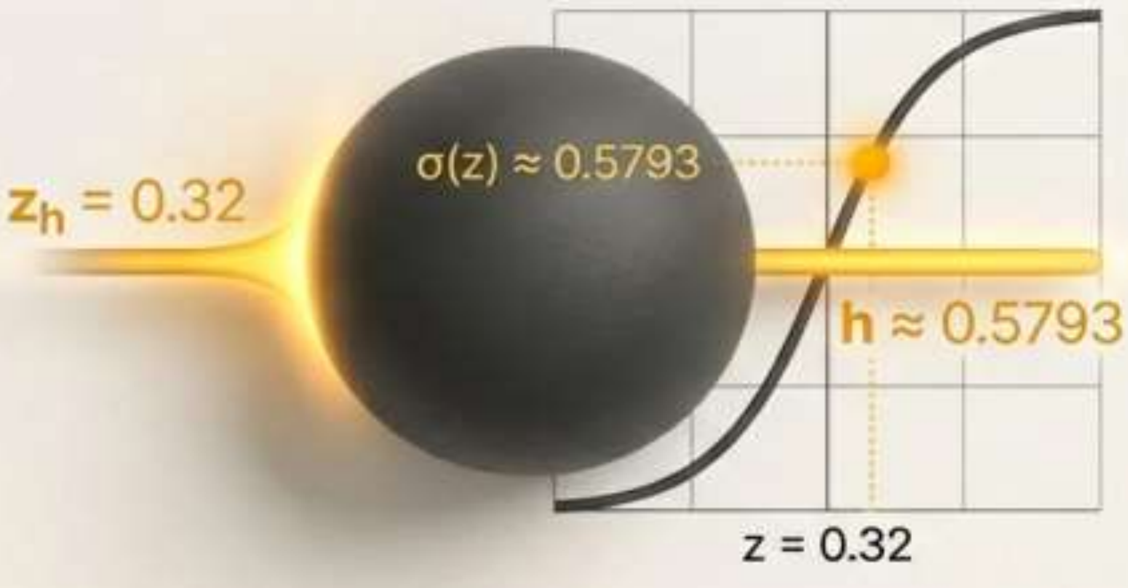
$$z_h = (0.6 \times 0.5) + (0.2 \times -0.4) + 0.1$$

$$z_h = 0.30 - 0.08 + 0.1$$

$$\mathbf{z_h = 0.32}$$

The neuron's net input, z_h, is the weighted sum of its inputs plus its bias.

The Neuron's Decision: Applying the Activation Function



Activation Function

Sigmoid $\sigma(z) = \dfrac{1}{1 + e^{-z}}$

Calculation

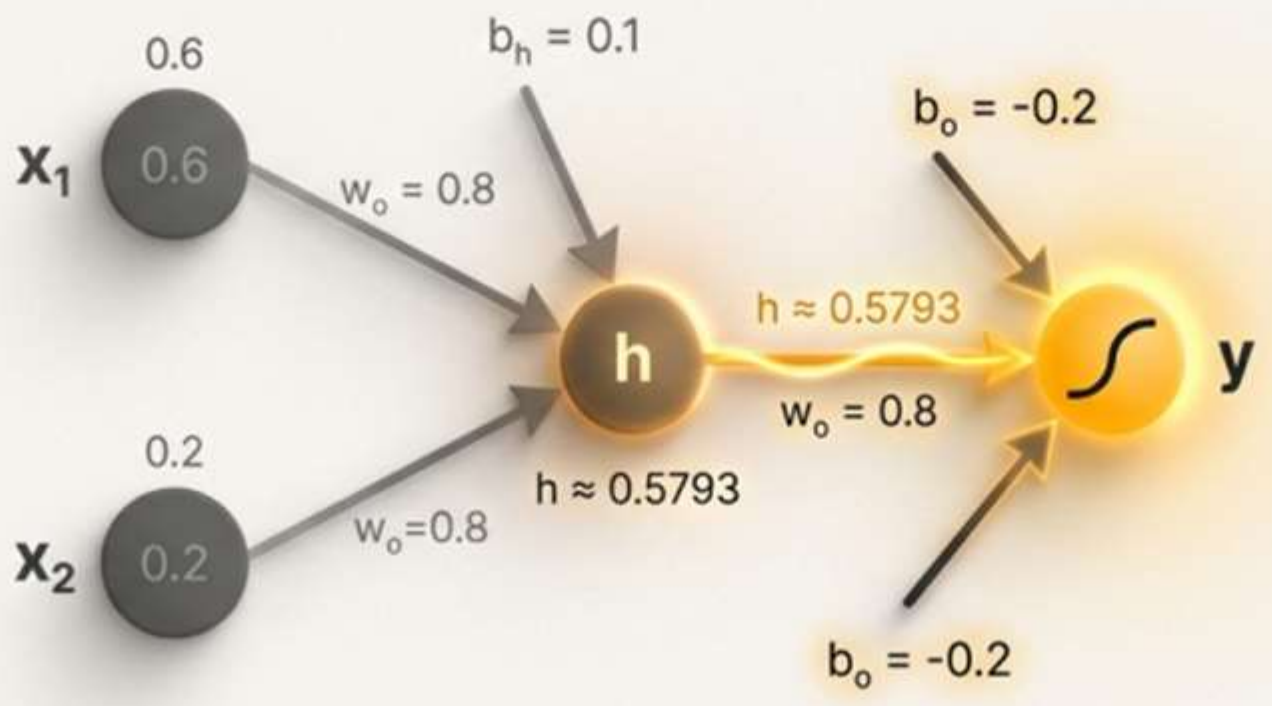
$h = \sigma(z_h) = \sigma(0.32)$

$h = \dfrac{1}{1 + e^{-0.32}}$

$h \approx 0.5793$

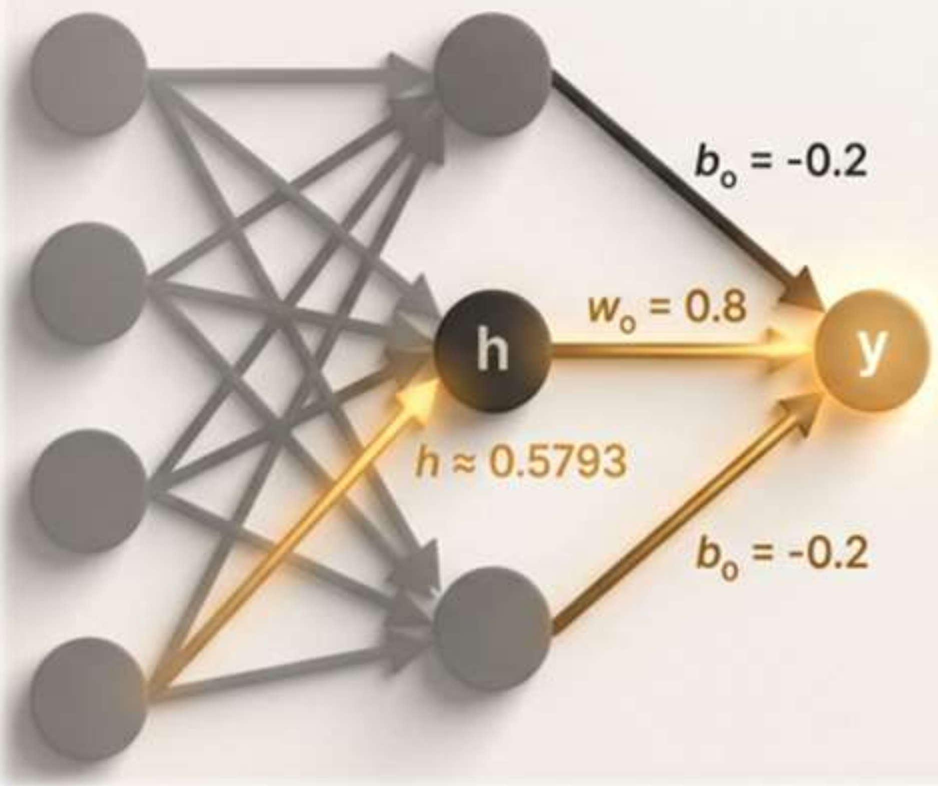
The sigmoid function transforms the input, scaling the output between 0 and 1. This is the new signal we send forward.

The Signal Advances to the Output Neuron



The hidden neuron's output, $h \approx 0.5793$, now travels along the connection to the final neuron, y. This signal will be weighted by w_o and combined with the output bias, b_o, in the next step.

Final Stop: The Output Neuron Aggregates its Signal



$z_o = (h \times w_o) + b_o$

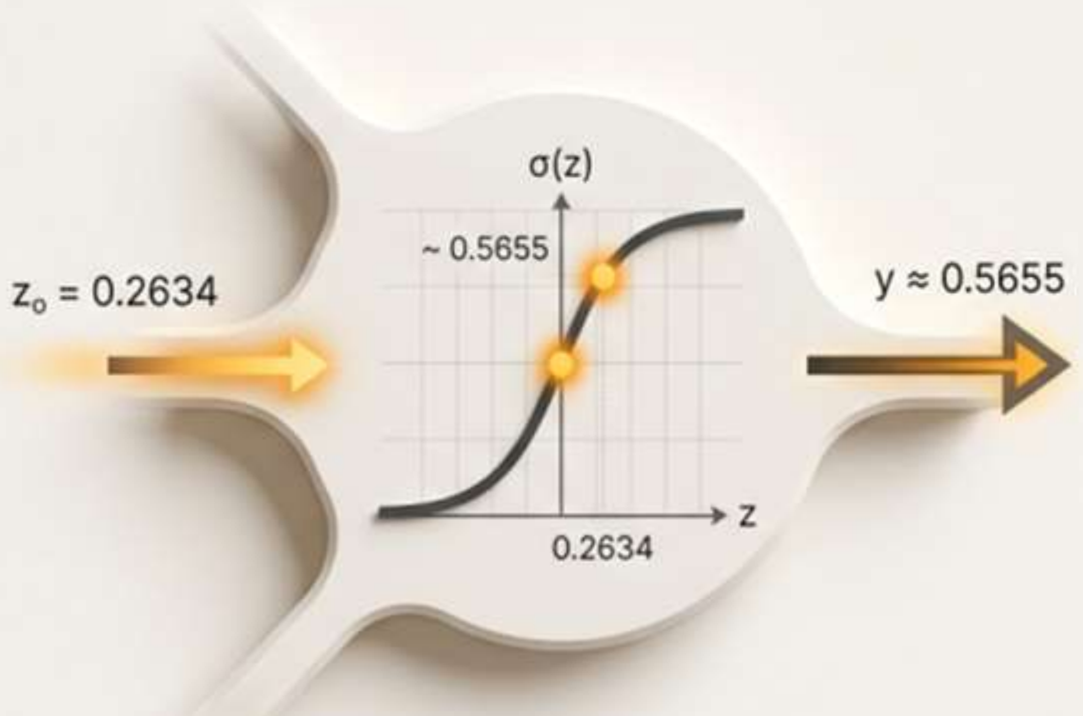
$z_o = (0.5793 \times 0.8) - 0.2$

$z_o = 0.4634 - 0.2$

$\mathbf{z_o = 0.2634}$

The process repeats: the output neuron calculates its own net input from the signal it received.

The Final Transformation: Producing the Network's Output



Activation Function

Sigmoid $\sigma(z) = \dfrac{1}{1+e^{-z}}$

Calculation

$$y = \sigma(z_0) = \sigma(0.2634)$$

$$y = \frac{1}{1+e^{-0.2634}}$$

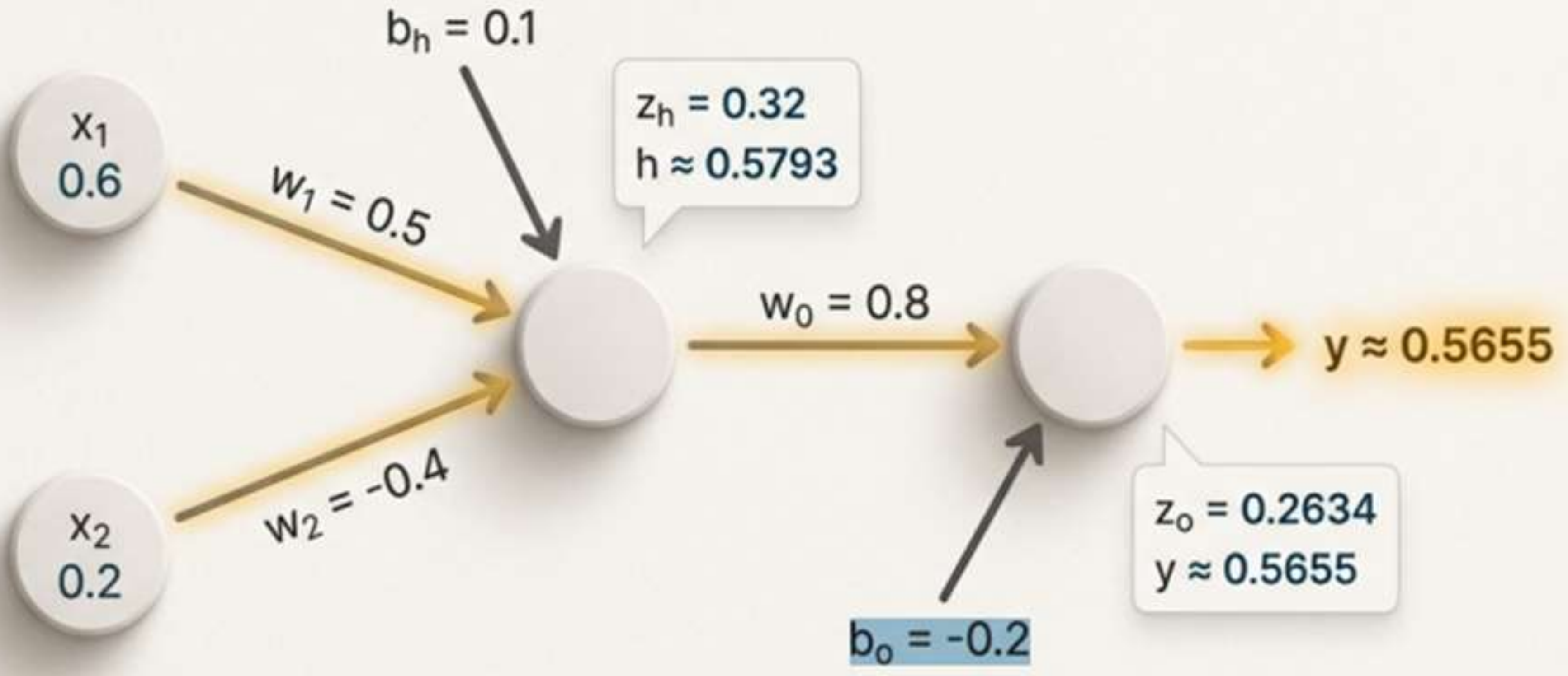
$y \approx 0.5655$

Key Takeaway

A final activation function gives us the network's definitive output.


Destination Reached
0.6
0.2
0.5655
Initial Inputs: [0.6, 0.2]
↓
Network Transformation
↓
Final Output: 0.5655
The network has processed the input signals
and produced a single output value.

The Complete Journey of the Signal



Take Home Assignment for Practice

- Consider a Single-layer feedforward neural network with the following.
- Inputs: $X_1 = 0.6$, $X_2 = 0.2$
- Weights (Input → Hidden): $w1 = 0.5$, $w2 = -0.4$
- Hidden bias (b_h) = 0.1
- Weight (Hidden → Output): $w_o = 0.8$
- Output bias (bo) = -0.2
- Activation: Sigmoid $\sigma(z) = 1 / (1 + e^{(-z)})$

Task needs to be done:

1. Net input to hidden neuron
2. Output of hidden neuron
3. Net input to output neuron
4. Final output of the network

Application: Solving the Software Telemetry Challenge

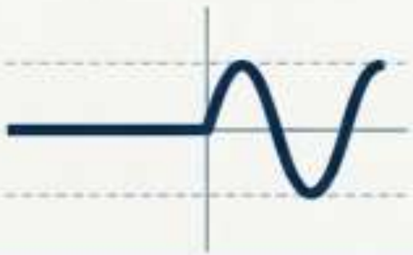
Recall our initial problem: traditional software telemetry struggles to classify the vast amounts of metric data collected by sensors. Fixed algorithms lack the ability to learn from this data.

"The main problem comes when it becomes very difficult to classify the collected metrics data & that's why if we need trained intelligent sensor based software project telemetry, i is ideal to use artificial neural network."

ANNs provide the missing 'learning factor," enabling the system to intelligently interpret complex patterns in metrics and improve the software development process.

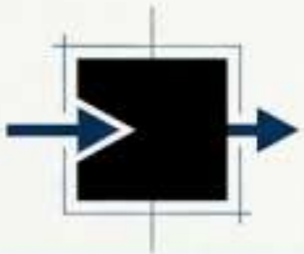
Capabilities of Neural Networks (Part 1)

A neural network derives its computing power from its massively parallel structure and its ability to learn and generalize. These properties provide good approximate solutions to complex problems that are otherwise intractable.



Nonlinearity

An essential property for processing real-world signals. The nonlinearity is *distributed* throughout the network.



Input-Output Mapping

Through *supervised learning*, the network learns from examples to construct a mapping from inputs to outputs, analogous to *nonparametric statistical inference.*

Adaptivity

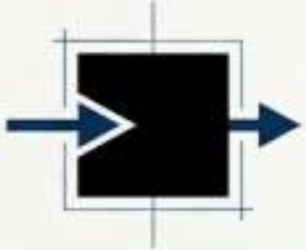
NNs adapt their weights to environmental changes, but must balance this with the **stability-plasticity dilemma** to avoid being too rigid or too sensitive.

Capabilities of Neural Networks (Part 1)

A neural network derives its computing power from its massively parallel structure and its ability to learn and generalize. These properties provide good approximate solutions to complex problems that are otherwise intractable.

Nonlinearity

An essential property for processing real-world signals. The nonlinearity is *distributed* throughout the network.



Input-Output Mapping

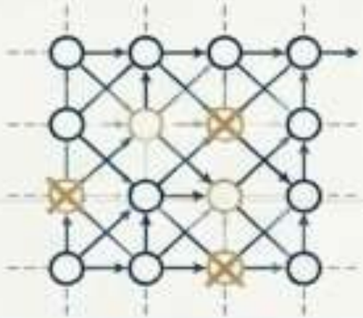
Through *supervised learning*, the network learns from examples to construct a mapping from inputs to outputs, analogous to *nonparametric statistical inference*.

Adaptivity

NNs adapt their weights to environmental changes, but must balance this with the **stability-plasticity dilemma** to avoid being too rigid or too sensitive.

Capabilities of Neural Networks (Part 2)

Building upon the foundational properties, these capabilities further highlight the unique advantages and operational characteristics of neural networks in handling complex tasks.

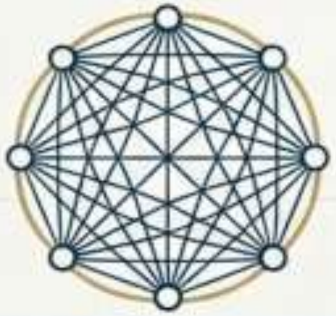


Fault Tolerance

Due to distributed information storage, the network's performance *degrades gracefully* rather than failing catastrophically if some neurons are damaged.

Evidential Response

A network can provide not only a decision but also a measure of *confidence*, allowing it to reject ambiguous patterns.

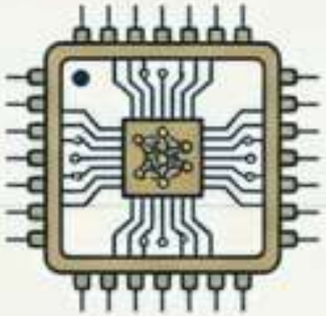


Contextual Information

Knowledge is represented by the entire network structure. Every neuron is potentially affected by the global activity of all other neurons.

Capabilities of Neural Networks (Part 3)

Building upon the foundational properties, these capabilities further highlight the unique advantages and operational characteristics of neural networks in handling complex tasks.

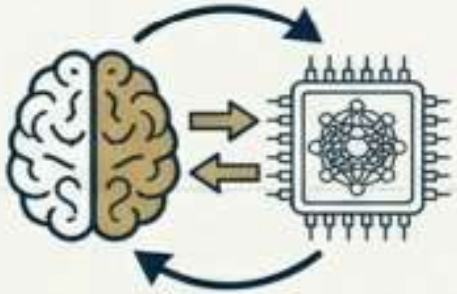


VLSI Implementability

The massively parallel nature of NNs makes them well-suited for implementation in Very-Large-Scale-Integrated (VLSI) technology.

Uniformity of Analysis and Design

Theories and learning algorithms can be shared across different applications, enabling the seamless integration of modular networks.



Neurobiological Analogy

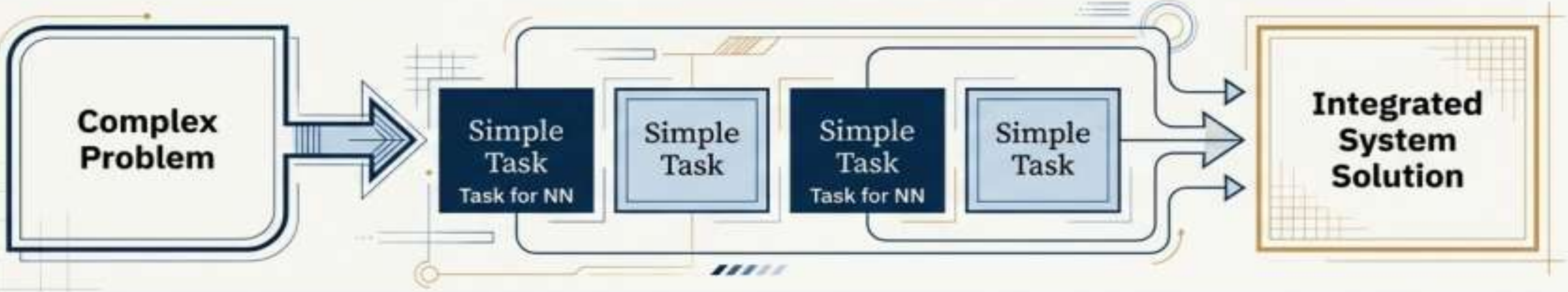
A two-way street: engineers gain inspiration for new designs, while neurobiologists use NNs as tools to understand the brain.

A Powerful Tool, Not a Panacea

In practice, neural networks do not solve problems by working individually. They must be integrated into a consistent system engineering approach.

The Process

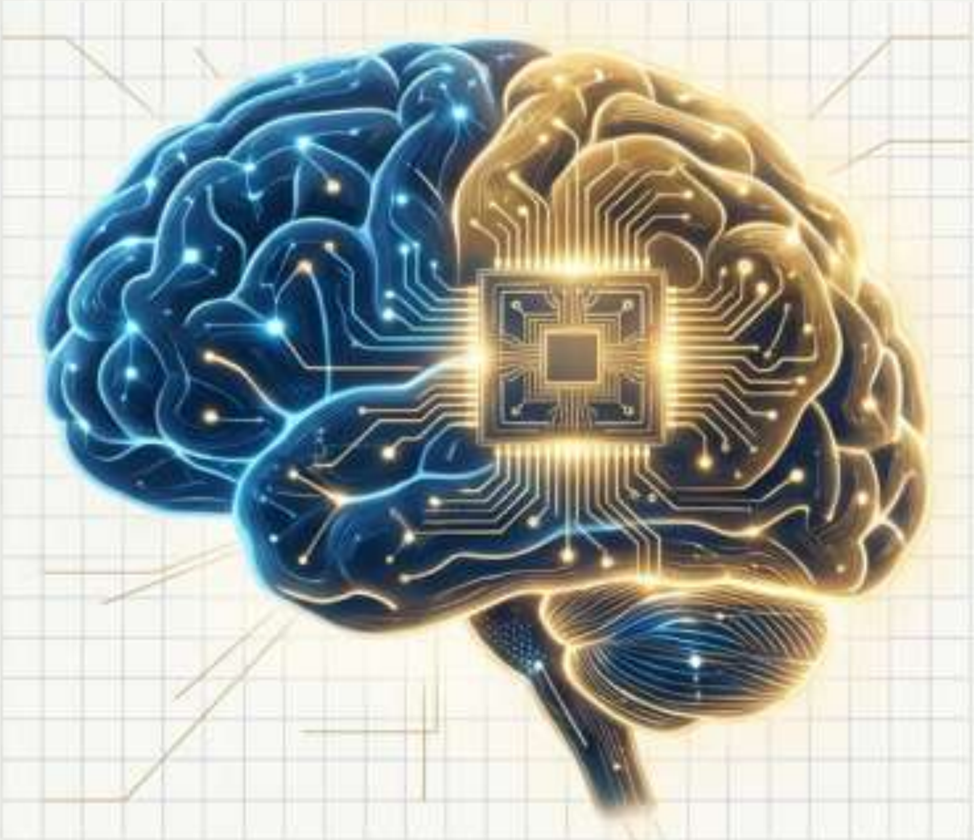
1. "A complex problem of interest is decomposed into a number of relatively simple tasks..."
2. "...and neural networks are assigned a subset of the tasks that match their inherent capabilities."



The Quest for Understanding Continues

A Statement of Humility

Source Serif Pro Regular
"*The artificial neurons we use... are truly primitive in comparison with those found in the brain. The neural networks we are presently able to design are just as primitive compared with the local circuits... in the brain.*"



A Concluding Vision

"However, the progress on so many fronts is remarkable. *With neurobiological analogy as the source of inspiration, and the wealth of theoretical and computational tools we are bringing together, it is certain that our understanding of artificial neural networks and their applications will continue to grow in depth as well as breadth, year after year.*"

Topics Coved:

- ✓ Introduction, Motivation and History,
- ✓ Components of a Neuron-synapses, dendrite, cell nucleus, axon.
- ✓ Important Terminologies of ANNs: Propagation function, Activation function, output function,
- ✓ Components of Artificial Neural Network: common activation functions, network topologies- feed forward, recurrent networks, completely linked networks

Thanks for your Attention!

You Can Post your doubts:

https://www.gcjana.in/courses/shardauniversity/2501/CSE472/#Post-doubts

Course Page:
https://www.gcjana.in/courses/shardauniversity/2501/CSE472/