Fundamentals of Artificial Neural Networks (I)

AI: Deep Learning and Neural Networks

Thuan L Nguyen, PhD

Slide 2: Fundamentals of Artificial Neural Networks

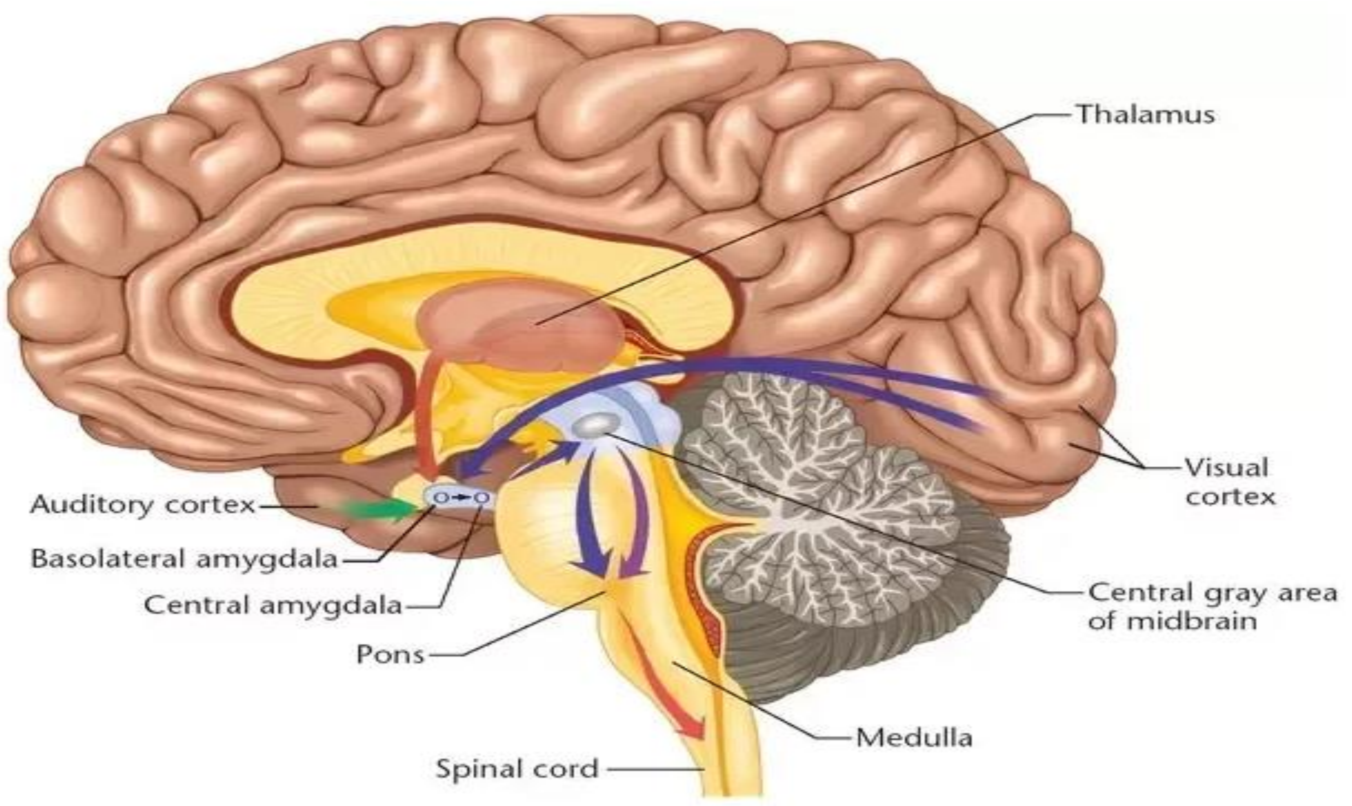
AI Deep learning (Source: mindovermachines.com)

Slide 3: Fundamentals of Artificial Neural Networks

1. Biological Neurons and Neural Networks
2. Artificial Neurons and Perceptron
3. Perceptron: A Simple Neural Network
4. Artificial Neural Networks: An Introduction
5. Artificial Neural Networks: Computation Power
6. Artificial Neural Networks: Architectures
7. Artificial Neural Networks: Applications

Slide 4: Fundamentals of Artificial Neural Networks

Biological Neural Networks: Human Brain



Human Brain (Source: Quora.com)

Slide 5: Fundamentals of Artificial Neural Networks

Biological Neural Networks

- The brain is one of the largest and most complex organs in the human body.

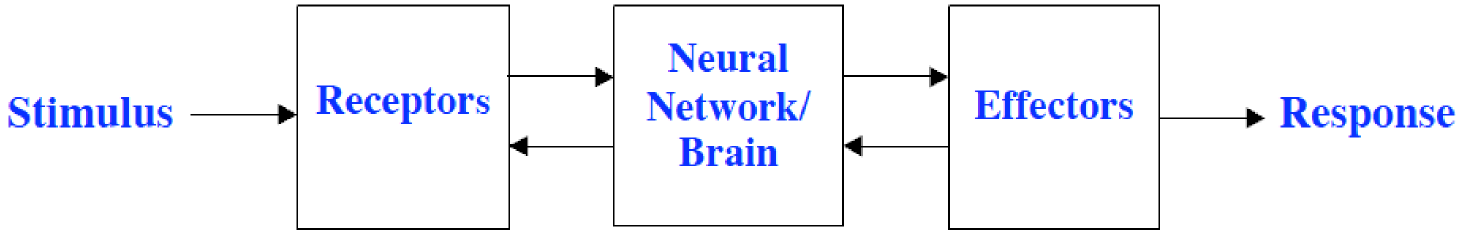
- It is made up of more than 100 billion nerves that communicate in trillions of connections called synapses.

- The brain is made up of many specialized areas that work together:
  - The cortex is the outermost layer of brain cells. Thinking and voluntary movements begin in the cortex.
  - The brain stem is between the spinal cord and the rest of the brain. Basic functions like breathing and sleep are controlled here.
  - The basal ganglia are a cluster of structures in the center of the brain. The basal ganglia coordinate messages between multiple other brain areas.
  - The cerebellum is at the base and the back of the brain. The cerebellum is responsible for coordination and balance.

Slide 6: Fundamentals of Artificial Neural Networks

Biological Neural Networks: Neural System

The human nervous system can be broken down into three stages that may be represented in block diagram form as:



Human Brain: Stimulus & Response (Source: Wikipedia)

- The receptors collect information from the environment – e.g. photons on the retina.
- The effectors generate interactions with the environment – e.g. activate muscles.
- The flow of information/activation is represented by arrows – feedforward and feedback.
- Naturally, this module will be primarily concerned with how the neural network in the middle works, but understanding its inputs and outputs is also important.

Slide 7: Fundamentals of Artificial Neural Networks

Biological Neural Networks: Levels of Brain Organization

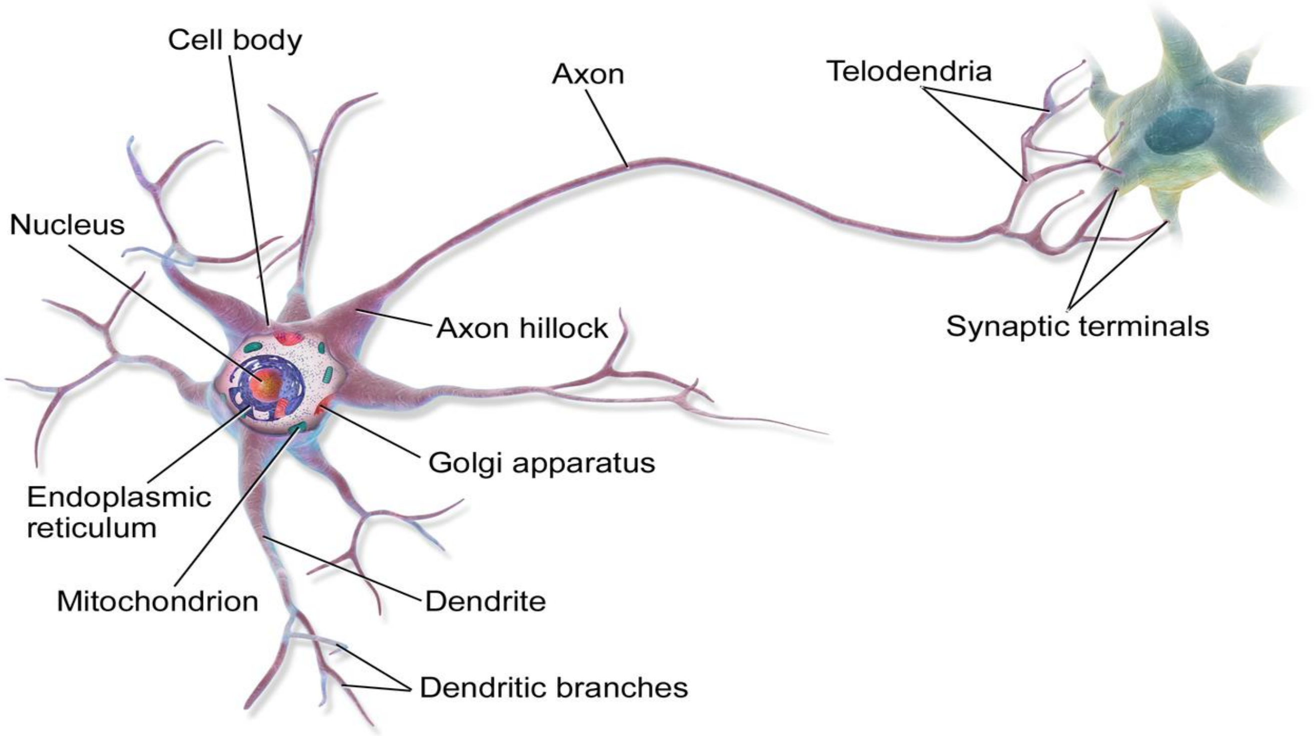
- The brain contains both large scale and small scale anatomical structures and different functions take place at the higher and lower levels.

- There is a hierarchy of interwoven levels of organization:
  1. Molecules and Ions
  2. Synapses
  3. Neuronal microcircuits
  4. Dendritic trees
  5. **Neurons**
  6. **Local circuits**
  7. Inter-regional circuits
  8. Central nervous system

- The artificial neural networks studied in this module are mostly approximations of levels 5 and 6.

Slide 8: Fundamentals of Artificial Neural Networks

Biological Neural Networks: Basic Components of Biological Neurons



Human Neuron (Source: by Bruce Blaus, is licensed under CC BY 3.0)

Slide 9: Fundamentals of Artificial Neural Networks

Biological Neural Networks: Basic Components of Biological Neurons

- The majority of neurons encode their **activations** or **outputs** as a series of **brief electrical pulses** (i.e. **spikes** or **action potentials**).

- The neuron's cell body (soma) processes the incoming activations and converts them into output activations.

- The neuron's nucleus contains the genetic material in the form of DNA. This exists in most types of cells, not just neurons.

- **Dendrites** are fibers which emanate from the cell body and provide the receptive zones that receive activation from other neurons.

- **Axons** are fibers acting as transmission lines that send activation to other neurons.

- The **junctions** that allow signal transmission between the axons and dendrites are called **synapses**. The process of transmission is by diffusion of chemicals called **neurotransmitters** across the synaptic cleft.

Slide 10: Fundamentals of Artificial Neural Networks

Biological Neural Networks: Neural Signal Processing

- **Signals** from connected neurons are collected by the **dendrites**.
- The **cells body** (soma) **sums** the incoming signals (spatially and temporally).
- When **sufficient** input is received (i.e., a threshold is exceeded), the neuron generates an **action** potential or '**spike**' (i.e., it 'fires').
- That **action potential** is transmitted along the axon to other neurons, or to structures outside the nervous systems (e.g., muscles).
- If sufficient input is **not** received (i.e., the threshold is **not** exceeded), the inputs quickly decay and **no** action potential is generated.
- **Timing** is clearly important – input signals must arrive together. Strong inputs will generate more action potentials per unit time.

Slide 11: Fundamentals of Artificial Neural Networks

Biological Neural Networks: Rate Coding vs. Spike Coding

- In biological neural networks, the individual spike timings are often important. So "**spike time coding**" is the most realistic representation for artificial neural networks.

- However, **averages** of spike rates across time or populations of neurons carry a lot of the useful information, and so "**rate coding**" is a useful approximation.

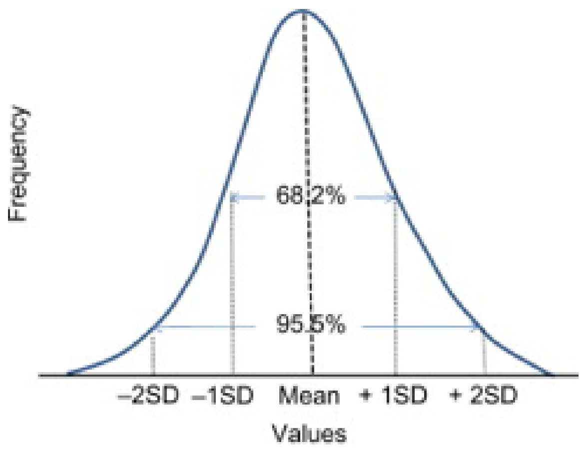
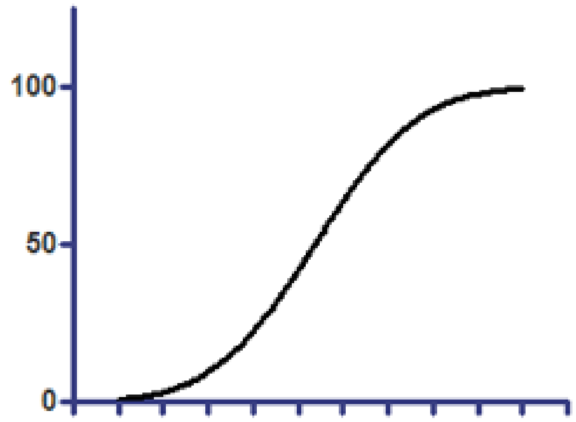
- Spike coding is more powerful, but the computer models are much more complicated and more difficult to train.

- Rate coding blurs the information coded in individual neurons, but usually leads to simpler models with differentiable outputs, which we will see later is important for generating efficient learning algorithms.

- Sigmoid shaped activation functions in the **rate coding** approach follow from the cumulative effect of Gaussian distributed spikes.

Slide 12: Fundamentals of Artificial Neural Networks

Biological Neural Networks: Rate Coding vs. Spike Coding

Gaussian/Normal Distribution	Cumulative Gaussian Distribution
	

Gaussian and Cumulative Gaussian Distribution (Source: Wikipedia)

Slide 13: Fundamentals of Artificial Neural Networks

Biological Neural Networks: Learning Process

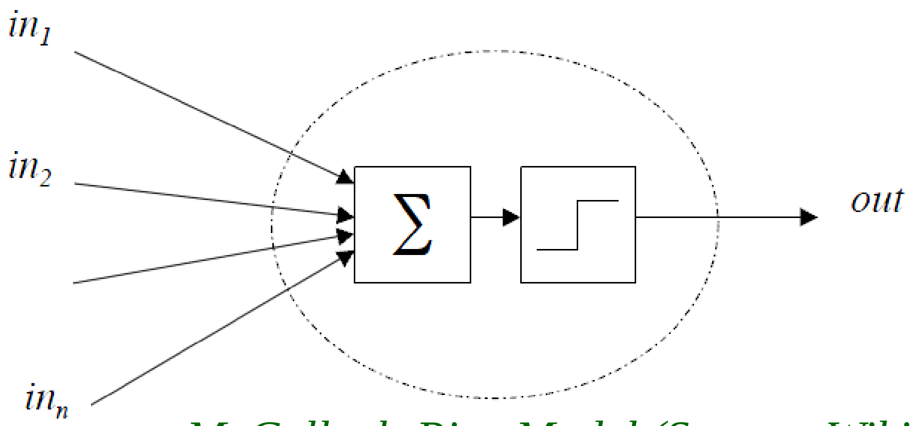
- Learning by iterative improvement:
  - Start with an initial (possibly random) solution.
  - Then improve on the solution step-by-step.

- Genetic Learning:
  - Based on evolution and natural selection.
    - ‘Evolve’ new solutions from old ones
    - Then ‘selection’ the new solutions which are good.

Slide 14: Fundamentals of Artificial Neural Networks

Artificial Neurons: The McCulloch-Pitts Neuron

A **simple rate coding model** of real neurons is also known as a **Threshold Logic Unit** :



McCulloch-Pitts Model (Source: Wikipedia)

- A set of synapses (i.e. connections) brings in activations, i.e., **inputs**, from other neurons.
- A **processing unit** sums the inputs, and then applies a **non-linear activation function**
  - Is also often called a threshold or transfer or squashing function
- An **output** line transmits the result to other neurons.

Slide 15: Fundamentals of Artificial Neural Networks

Artificial Neurons: The McCulloch-Pitts Neuron: Neuron Equation

- 1943: McCulloch and Pitts proposed the McCulloch-Pitts neuron model

We can now write down the equation for the output Y_j of a McCulloch-Pitts neuron as a function of its inputs I_i:

$$Y_j = \mathrm{sgn}\left(\sum_{i=1}^{n} I_i - \theta\right)$$

where θ is the neuron's **activation threshold**. When

$$Y_j = 1, \quad if \sum_{k=1}^{n} I_k \geq \theta \qquad\qquad Y_j = 0, \quad if \sum_{k=1}^{n} I_k < \theta$$

Slide 16: Fundamentals of Artificial Neural Networks

Artificial Neurons: The McCulloch-Pitts Neuron

- In mathematics, the sign function or signum function (from signum, Latin for "sign") is an odd mathematical function that extracts the sign of a real number. In mathematical expressions the sign function is often represented as **sgn**.

The signum function of a real number x is defined as follows:

$$\operatorname{sgn}(x) := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

Alternatively:

$$\operatorname{sgn}(x) = \frac{\mathrm{d}}{\mathrm{d}x}|x|, \quad x \neq 0$$

Slide 17: Fundamentals of Artificial Neural Networks

Perceptron: The Simplest Form of a Neural Network

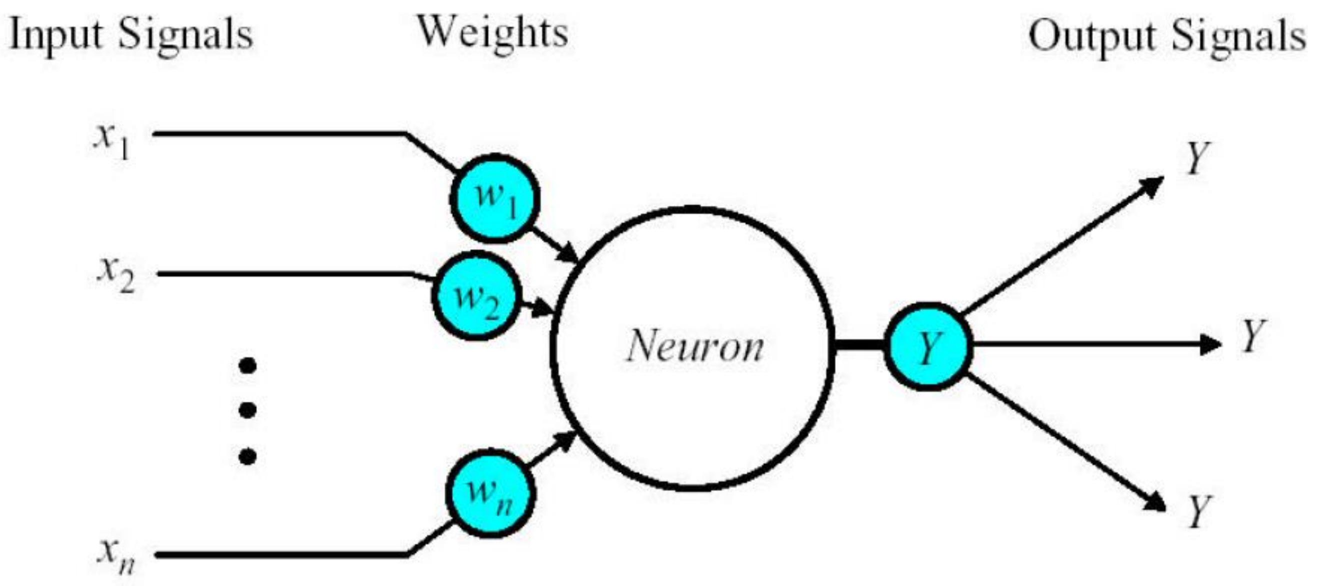
- A perceptron:
  - Frank Rosenblatt introduced the concept of a perceptron (1958):
    - He proposed a training algorithm that provided the first procedure for training a simple artificial neural network called perceptron.

- Perceptron: The simplest form of a neural network.
  - It consists of a single neuron with adjustable synaptic weights and a hard limiter.

Slide 18: Fundamentals of Artificial Neural Networks

Perceptron: A Network of The McCulloch-Pitts Neurons

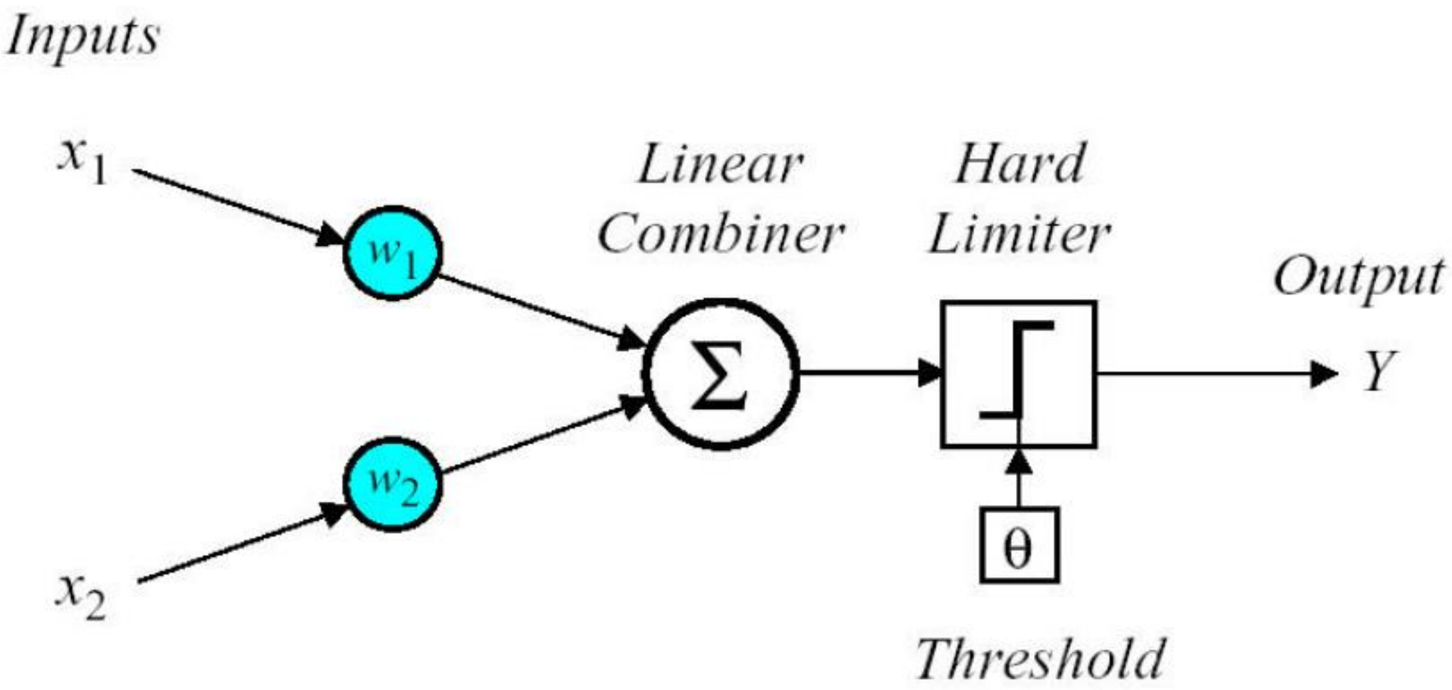
A **simple rate coding model** of real neurons is also known as a **Threshold Logic Unit** :



McCulloch-Pitts Model (Source: Wikipedia)

Slide 19: Fundamentals of Artificial Neural Networks

Perceptron: A Network of The McCulloch-Pitts Neurons



McCulloch-Pitts Model (Source: Wikipedia)

Slide 20: Fundamentals of Artificial Neural Networks

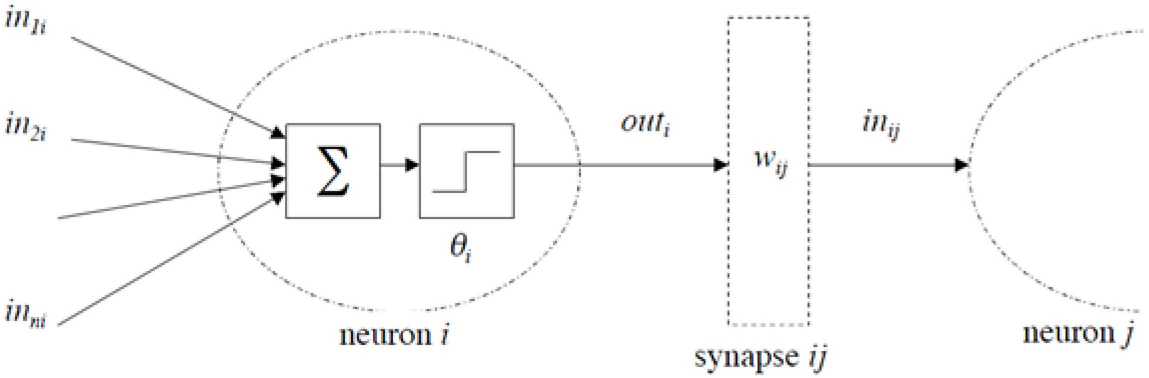
Perceptron: A Network of The McCulloch-Pitts Neurons

- Frank Rosenblatt introduced the concept of a perceptron (1958)
  - Each input I_i is multiplied by a weight w_{ji} (synaptic strength)
  - These weighted inputs are summed to give the activation level, A_j
  - The activation level is then transformed by an activation function to produce the neuron's output, Y_i
  - W_{ji} is known as the weight from unit i to unit j
    - $W_{ji} > 0$, synapse is excitatory
    - $W_{ji} < 0$, synapse is inhibitory
  - Note that I_i may be
    - External input
    - The output of some other neuron

Slide 21: Fundamentals of Artificial Neural Networks

Perceptron: Networks of McCulloh-Pitts Neurons

- To finish a meaningful computation task, it is necessary to have a network of multiple neurons:



$$out_k w_{ki} = in_{ki} \qquad out_i = step\left(\sum_{k=1}^{n} in_{ki} - \theta_i\right) \qquad out_i w_{ij} = in_{ij}$$

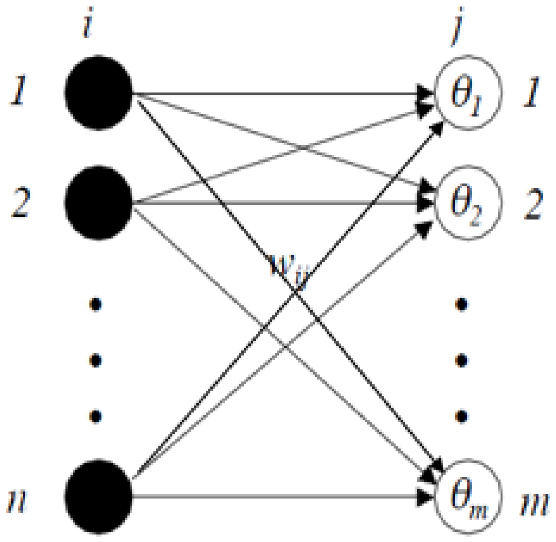
McCulloch-Pitts Model (Source: Wikipedia)

Slide 22: Fundamentals of Artificial Neural Networks

Deep Learning: Simple Single-Layer Neural Networks

Perceptron:

- The fundamental unit of an artificial neural network
- A simple – single-layer – artificial neural network:
  - A simple neural network that has one layer of input neurons feeding forward to one output layer of McCulloch-Pitts neurons, with full connectivity.



$$out_j = step\left(\sum_{i=1}^{n} in_i w_{ij} - \theta_j\right)$$

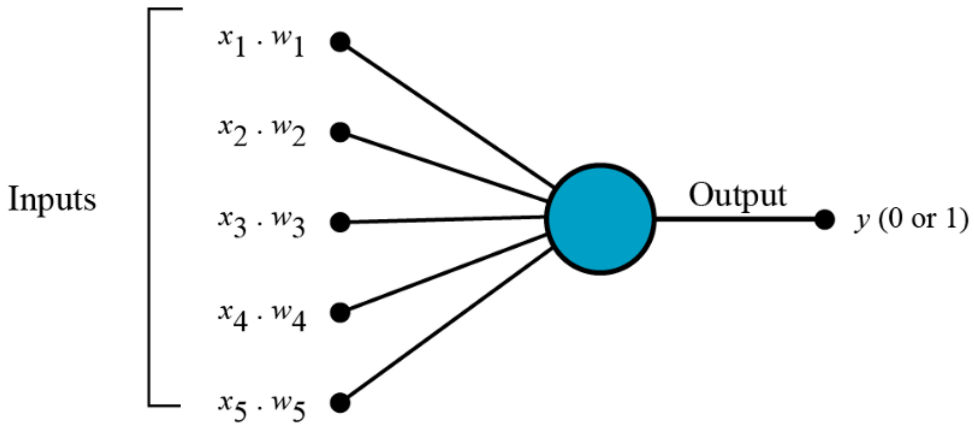
AI Deep Learning: Perceptron (Source: Wikipedia)

Slide 23: Fundamentals of Artificial Neural Networks

Deep Learning: Simple Single-Layer Neural Networks

Perceptron:

- The McCulloch-Pitts neuron model is actually the **simplest** single-layer neural network.
  - One or more inputs → One output
- Therefore, the McCulloch-Pitts neuron model represents a **perceptron,** the simplest neural network.

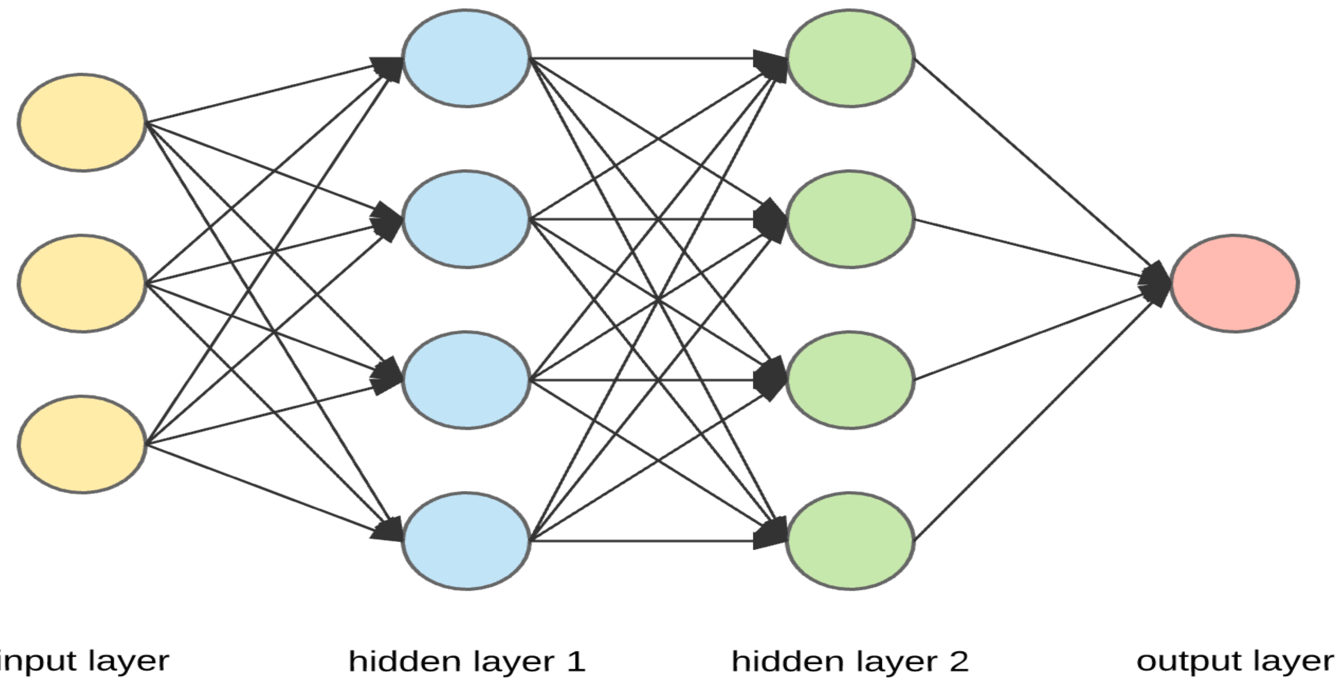


McCulloch-Pitts Model (Source: towardsdatascience.com)

Slide 24: Fundamentals of Artificial Neural Networks

Deep Learning: Multi-Layer Neural Networks

- Single-layer perceptrons: very limited regarding the computation power
- Multi-layer perceptrons, i.e., multi-layer neural networks, were constructed.

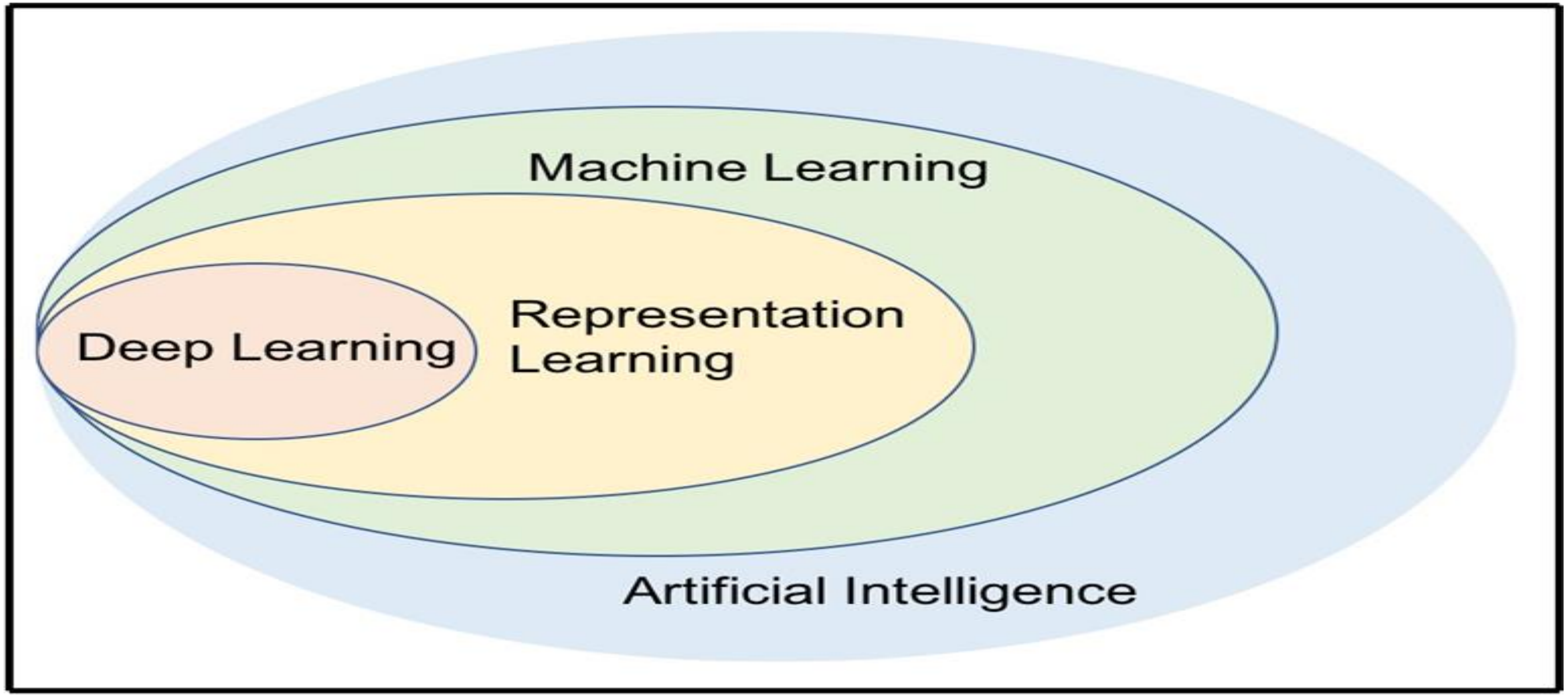


AI Deep Learning: Multi-layer Neural network (Source: medium.com)

Slide 17: AI Deep Learning: An Introduction

AI: Deep Learning: Neural Networks with Multiple Layers

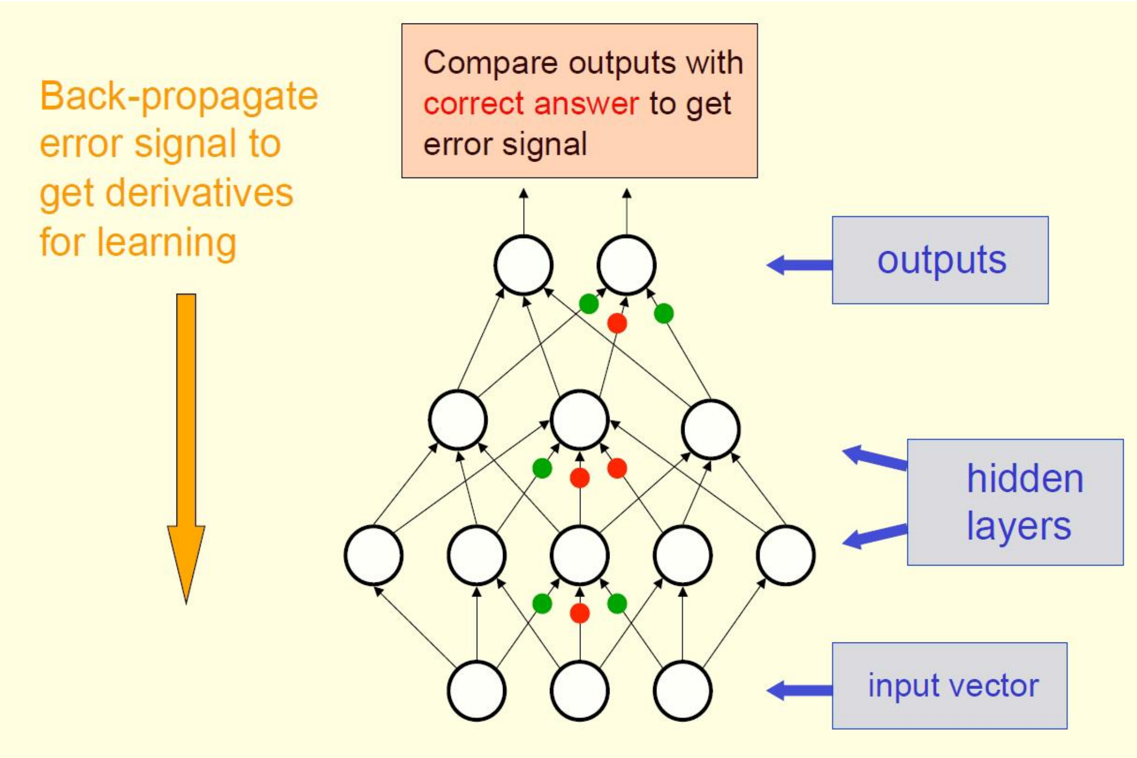
- Deep Learning and Other AI Approaches



Sources: Di, Bhardwaj, & Wei (2018)

Slide 12: AI Deep Learning: An Introduction

AI: Deep Learning: Neural Networks with Multiple Layers



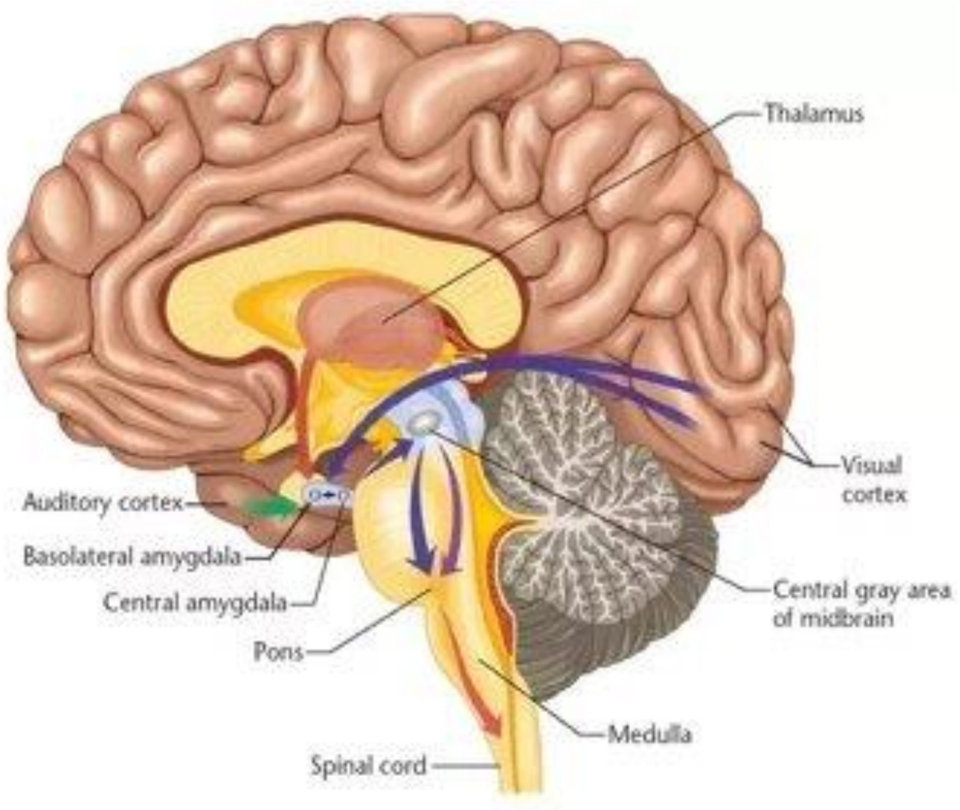
Deep learning (Sources: G. E. Hilton, 1997)

- "Deep Learning" stands for the concept of successive layers of representations.
  - How many layers contribute to a model of the data is called the **depth** of the model.
  - Other appropriate names for the field could have been **layered representations learning** and **hierarchical representations learning**.
  - **Modern deep learning** often involves tens or even hundreds of **successive layers** of **representations** that are all learned automatically from exposure to training data.

Slide 18: AI Deep Learning: An Introduction

AI Deep Learning: Multiple Layers: From a Biological Neural Viewpoint

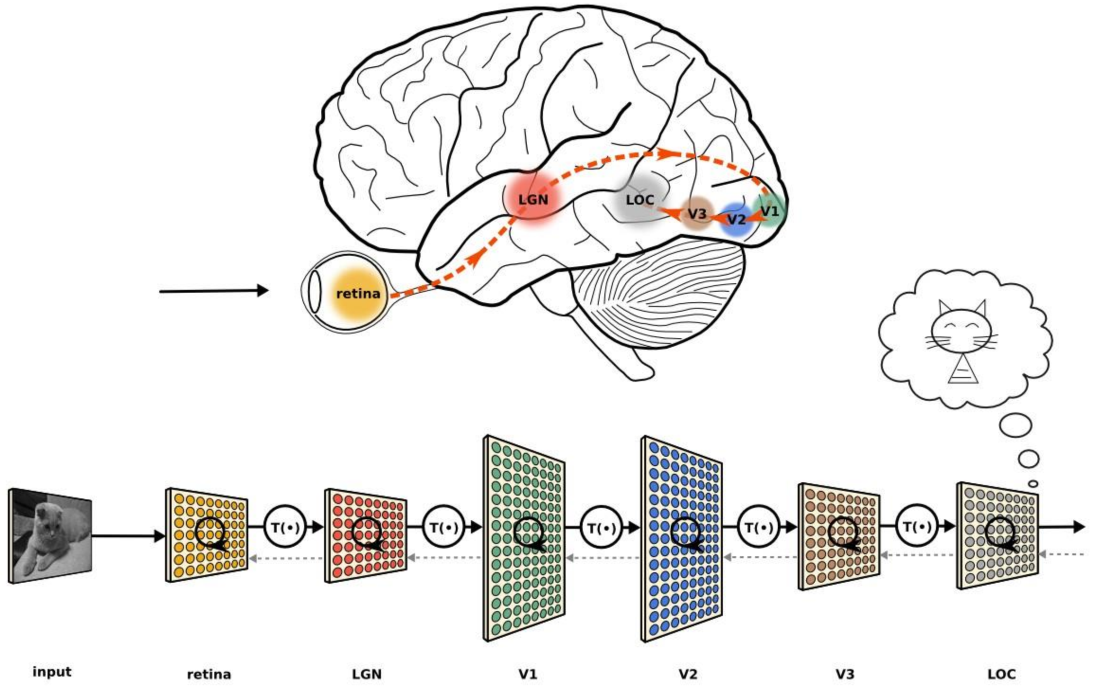
- An architecture for learning is biologically inspired.
- The human brain has deep architecture:
  - The cortex seems to have a generic learning approach.
- A given input is perceived at multiple levels of abstraction.
  - Each level corresponds to a different area of the cortex.
- We process information in hierarchical ways.
  - With multi-level transformation and representation.
- Therefore, we learn simple concepts first then compose them together.



Human Brain (Source: Quora.com)

Slide 19: AI Deep Learning: An Introduction

AI Deep Learning: Multiple Layers: From a Biologocal Neural Viewpoint



Biologocal Neural Viewpoint (Source: Wikipedia)

- The structure of understanding can be found in a human's vision system as shown in the figure:
  - Signal path from the retina to human lateral occipital cortex (LOC)
  - The path which finally recognizes the object
  - The ventral visual cortex comprises a set of areas that process images in increasingly more abstract ways, from edges, corners and contours, shapes, object parts to object
  - This path allows us to learn, recognize, and categorize three-dimensional objects from arbitrary two-dimensional views.

Slide 20: AI Deep Learning: An Introduction

AI Deep Learning: Multiple Layers: From a Representation Viewpoint

- For most traditional machine learning algorithms, their performance depends heavily on the representation of the data they are given.
  - Therefore, domain prior knowledge, feature engineering, and feature selection are critical to the performance of the output.
  - But hand- crafted features lack the flexibility of applying to different scenarios or application areas.
  - Also, they are not data-driven and cannot adapt to new data or information comes in.

- For many tasks related to various input formats such as **image**, **video**, **audio**, and **text**:
  - It is very difficult to know what kind of features should be extracted
    - Let alone their generalization ability for other tasks that are beyond the current application.
  - Manually designing features for a complex task requires a great deal of domain understanding, time, and effort.
    - Sometimes, it can take decades for a large group of researchers to make progress in this area.

Slide 21: AI Deep Learning: An Introduction

AI Deep Learning: Multiple Layers: From a Representation Viewpoint

- Representation Learning:
  - It is a data driven type of approach using machine learning to discover the representation.
    - Such representation can represent the mapping from representation to output (supervised), or simply representation itself (unsupervised).

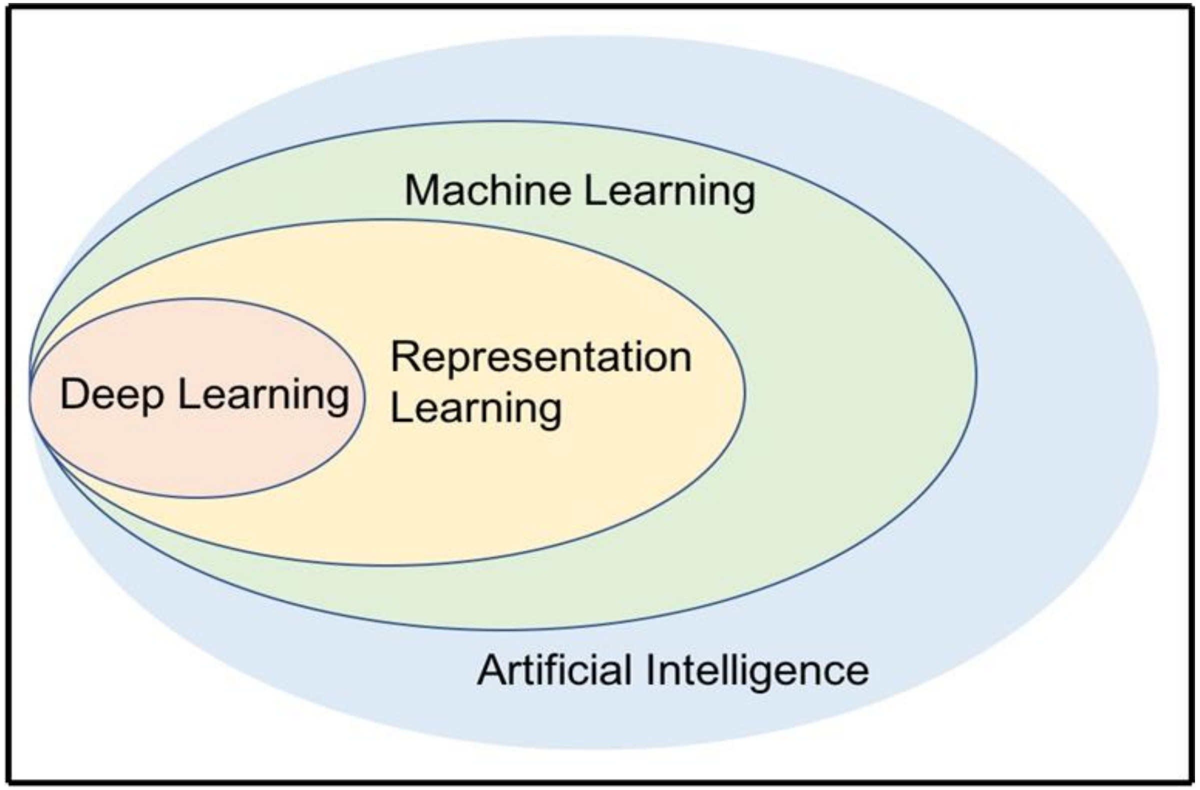
  - Learned representations often result in much better performance as compared to what can be obtained with hand-designed representations.
    - This also allows AI systems to rapidly adapt to new areas, without much human intervention.

  - With a representation learning algorithm, we can discover a good set of features for a simple task in minutes or a complex task in hours to months.
    - It may take vastly more time and effort if using hand-craft and design features.

Slide 22: AI Deep Learning: An Introduction

AI Deep Learning: Multiple Layers: From a Representation Viewpoint

- Deep Learning is Representation Learning
  - Deep learning feature extraction happens automatically when the deep architecture tries to process the data, learning, and understanding the mapping between the input and the output.
  - This brings significant improvements in accuracy and flexibility since human designed feature/feature extraction lacks accuracy and generalization ability.
  - In addition to this automated feature learning, the learned representations are both distributed and with a hierarchical structure.
  - Such successful training of intermediate representations helps feature sharing and abstraction across different tasks.



AI Deep Learning (Sources: Di, Bhardwaj, & Wei, 2018)