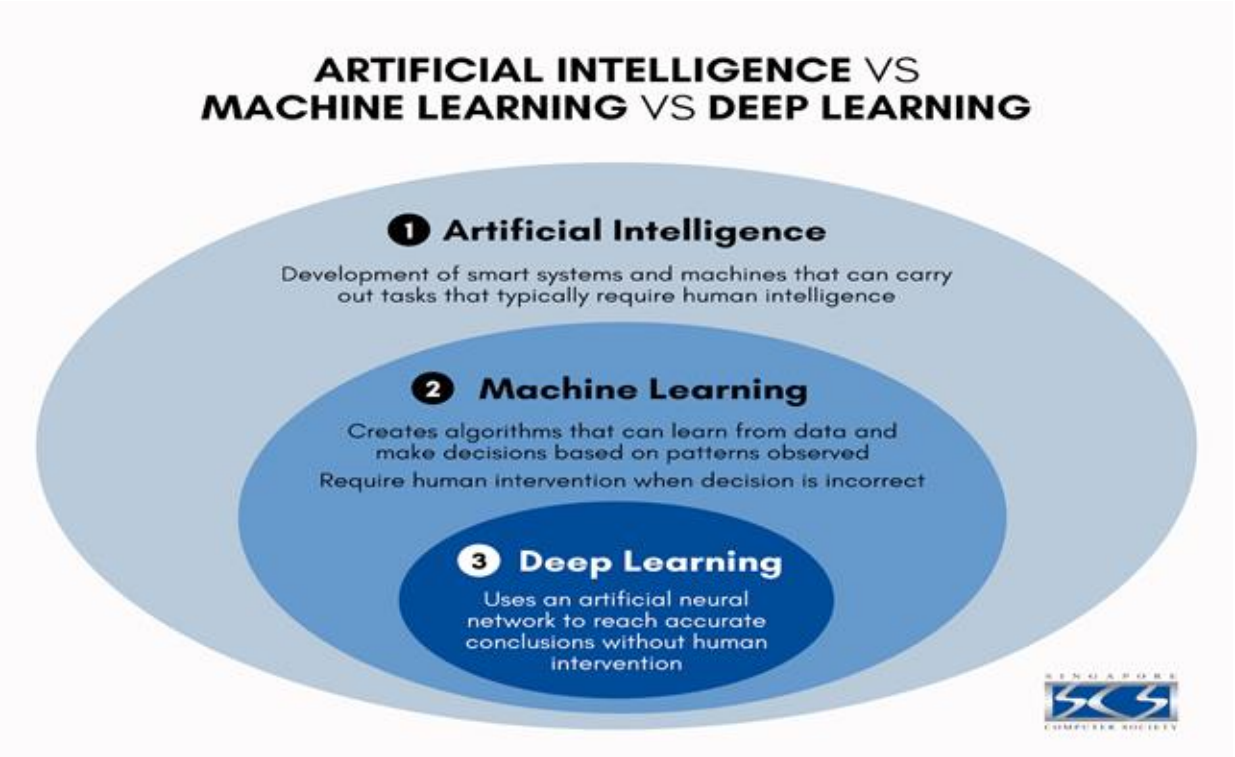
AI & Neural Networks

AI, Machine Learning, Deep Learning, Neural Networks, and the Math Behind Neural Networks

AI/ML/DL/Neural Networks? What is the relationship?



Presentation Structure

1. Different types of AIs (broadest)
2. ML and different types of ML
3. DL
4. Neural Networks (smallest)
    a. Math behind NNs
5. Applications of NNs and AIs

Artificial Intelligence

-machines that mimic human intelligence, actions, and cognitions like problem-solving and learning

Generally, an AI can…

1.Discover

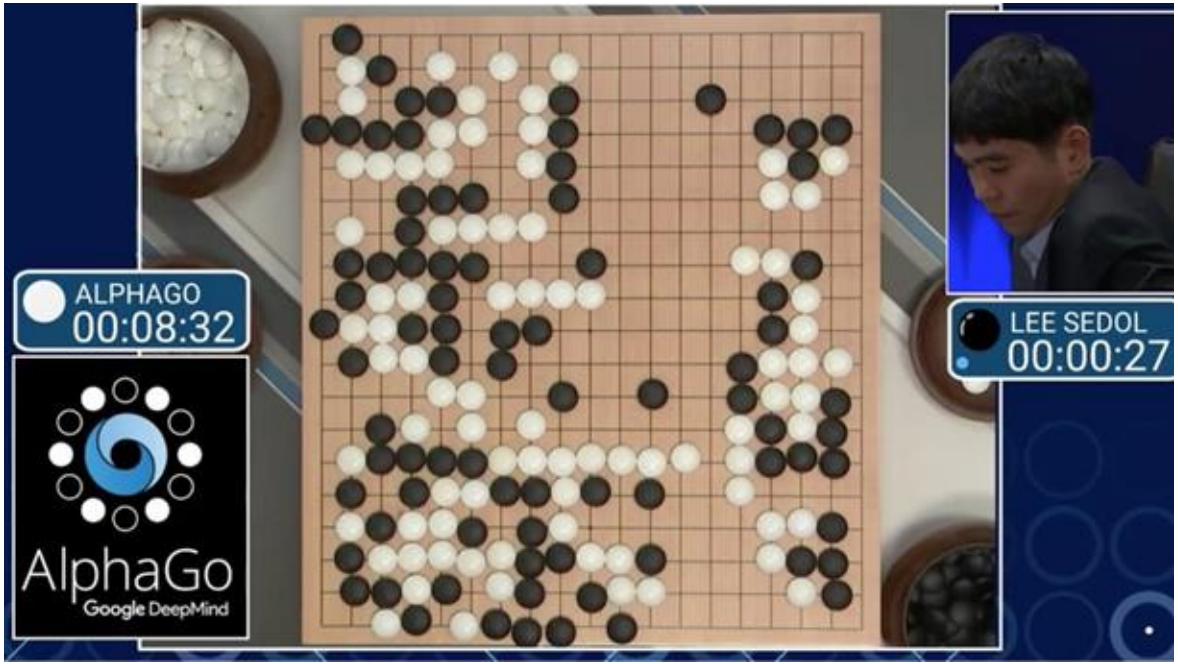
2.Infer

3.Reason (in a general sense)

Artificial Intelligence: Types of AI

- Artificial Narrow Intelligence (ANI)
- Artificial General Intelligence (AGI)
- Artificial Super Intelligence (ASI)

Artificial Intelligence: ANI (AlphaGo)



Artificial Intelligence: ANI (Autopilots)

Artificial Intelligence: ANI (Siri)

Artificial Intelligence: ANI (Generative AI like ChatGPT)

ChatGPT

Here's an illustration showcasing Siri as a friendly robot assistant, setting up reminders in a cozy, well-organized home office. The scene is filled with hints of Siri's dedication to helping with daily tasks.

Artificial Intelligence: AGI

-AGI: AI system that understands, learns, and applies its intelligence to solve any problem that a human being can, with the same efficiency or better

-AGIs adapt to new situations.

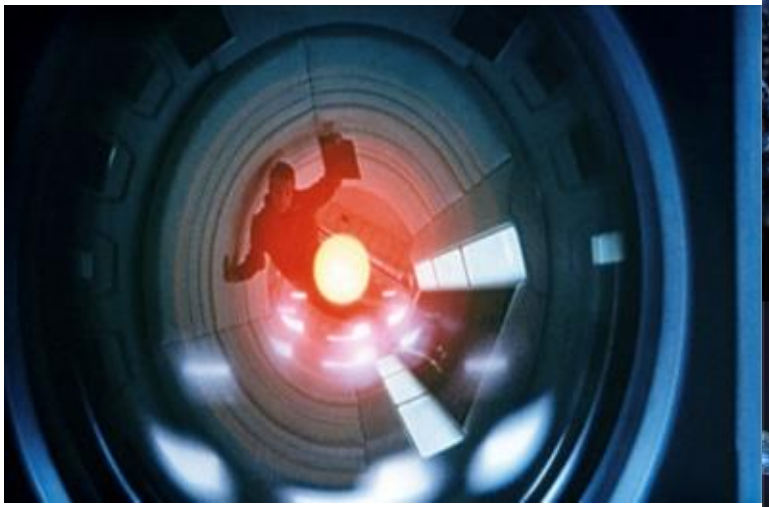
-Difference between ANI & AGI: scope of intelligence and adaptability.

-AGIs don’t currently exist.

Artificial Intelligence: ASI

-ASI: **outperforms the best human brains in practically every field**

-ASIs don’t currently exist.



Machine Learning: A Subset of AI

-Functionally, we may see Machine Learning as a way for computers to become AI/ANIs

-“the field of study that gives computers the ability to learn without explicitly being programmed.”

–Dr. Arthur Samuel, an AI pioneer

-Learning: the process by which a computer system improves its performance on a specific task over time, based on its experience with data

Machine Learning: Supervised Learning

-there is always a straightforward, yes-or-no answer about things the AI is trying to predict or describe.

-there is a human guide involved in the training, telling the algorithm what conclusions it should make regarding the yes-or-no question. .

Machine Learning: Unsupervised Learning

-no straightforward, yes-or-no answer

-about learning the underlying patterns and structures from the data

-e.g., customer segmentation: companies have no prior knowledge about the customers.

Deep Learning: A Subset of Machine Learning

Different from traditional ML in the following ways…

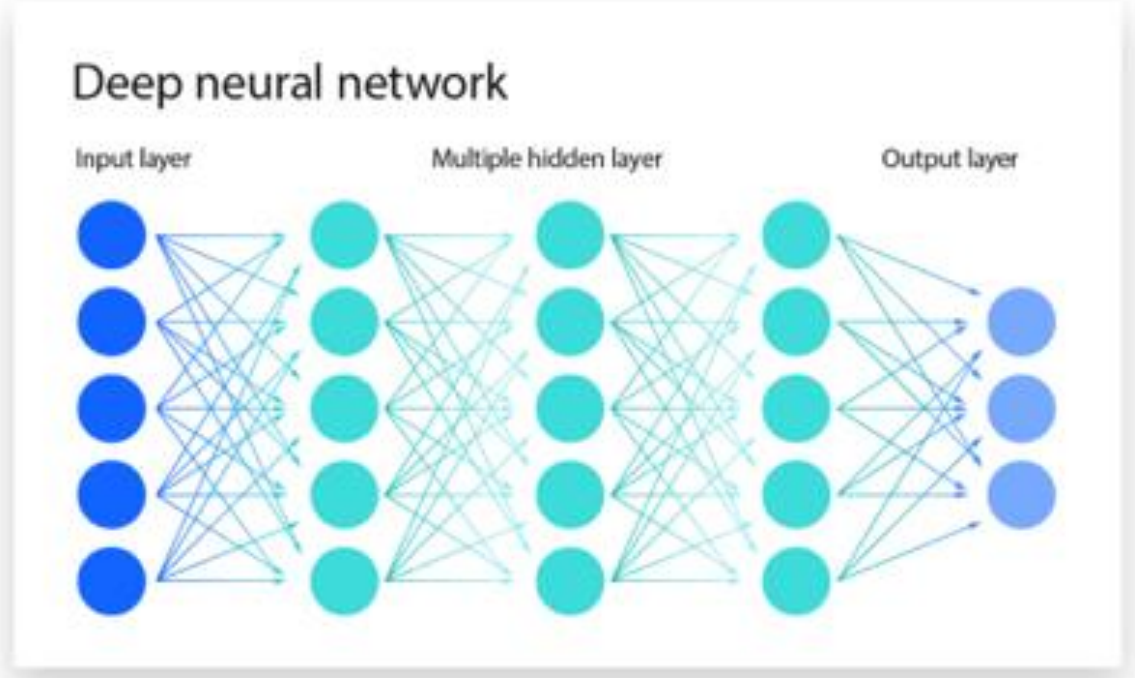
-Data Processing: eliminates some of manual data pre-processing that is typically involved with machine learning when dealing with inputs.

-Data type: relies on abundant and complex data

-Computational Power: requires more substantial computational resources like GPUs

-

Neural Networks: Architecture Used By DL



How Neural Networks Work?

1.Initialize the Network: how many layers, how many nodes in each layer; assign random values to weights and biases

2.Forward Propagation: feed training data into the network; through layer-by-layer transformation, you get a prediction output; calculating the loss function L

3.Backward Propagation (where learning happens): randomized weights and biases are updated with the intent to minimize the loss function; backward

4.Iterations: repeat step 1 through 3 until you meet a specific criterion

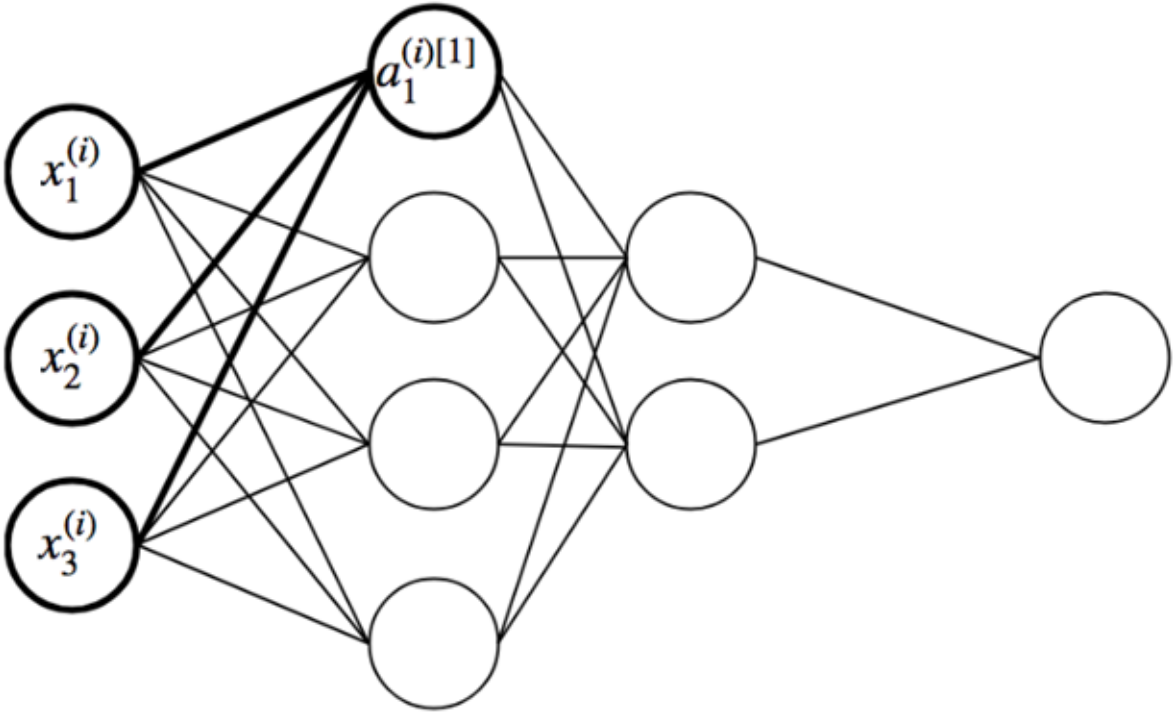
5.Validation and Testing

Math Behind Neural Networks

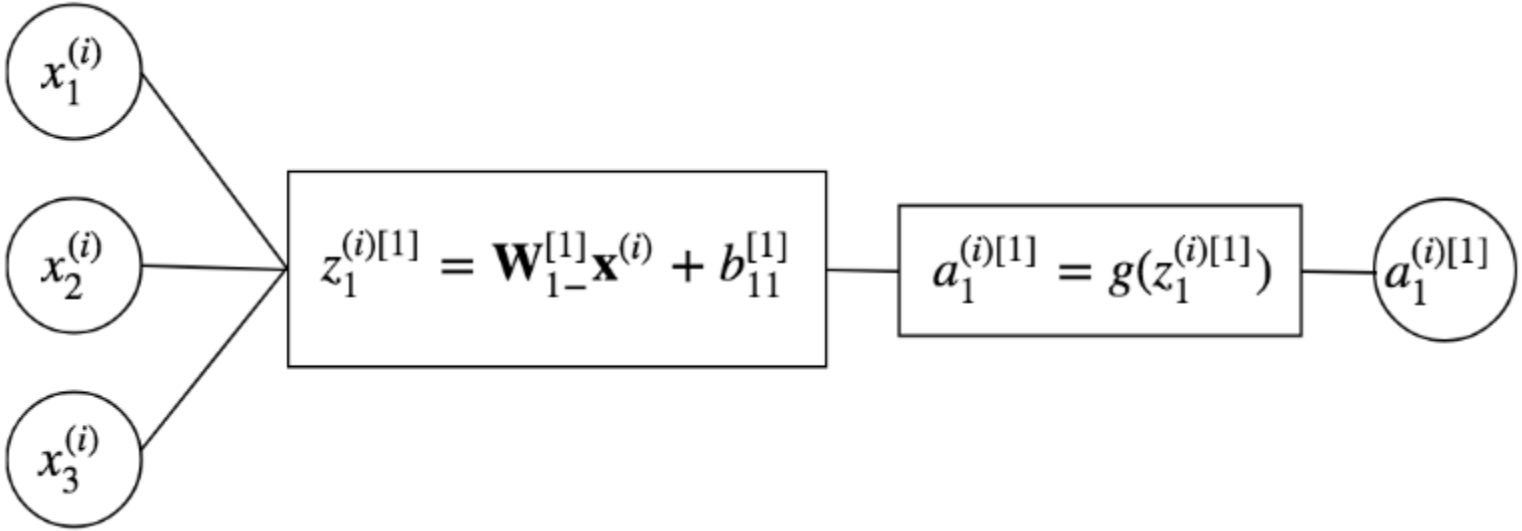
-Forward Propagation: Linear Algebra to represent the architecture of NNs

-Backward Propagation: Calculus to update the weights and biases to minimize the loss function

Forward Propagation: Architecture of Three Layer NN



Forward Propagation: Calculate alpha



Forward Propagation: Activation Function

e.g.,

$$g(z) = \begin{cases} z & \text{if } z > 0 \\ 0 & \text{otherwise} \end{cases}$$

The Loss Function: quantifying the difference between prediction & actual output

*Greater the loss function means greater difference and poorer prediction

*Minimize the loss function L in some way

Loss Calculation (Binary Cross-Entropy):

$$L = -(y \log(\hat{y}) + (1 - y) \log(1 - \hat{y}))$$

Backward Propagation: Chain Rule in Differential Calculus

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial a} \cdot \frac{\partial a}{\partial z} \cdot \frac{\partial z}{\partial w}$$

-dL/dw: from Calc 3 (gradient), we know this indicates the direction in which the loss function increases the fastest w.r.t the weight w, holding other vars constant.

-Since there are different weights/biases, like w_11, w_12, b_11,..., we calculate dL/dw_11, dL/dw_12, dL/db_11 in practice..

Backward Propagation: Gradient Descent (A Weight Update Rule)

$$w_{\text{new}} = w_{\text{old}} - \eta \cdot \frac{\partial L}{\partial w}$$

-Mathematically, this is how we update the weights/biases

-Given Calc 3 result (dL/dw is the direction of max. increase), -dL/dw is the direction in which the loss function decreases the fastest

-Update weights (or biases): calculating dL/dw (or dL/db) and manually giving eta (hyperparameter: learning rate) a value

Backward Propagation: Different Types of Gradient Descent

1. **Batch Gradient Descent (BGD)**: Use the entire training dataset and weights/biases updated using a mean gradient

2. **Stochastic Gradient Descent (SGD)**: Use training sample one by one

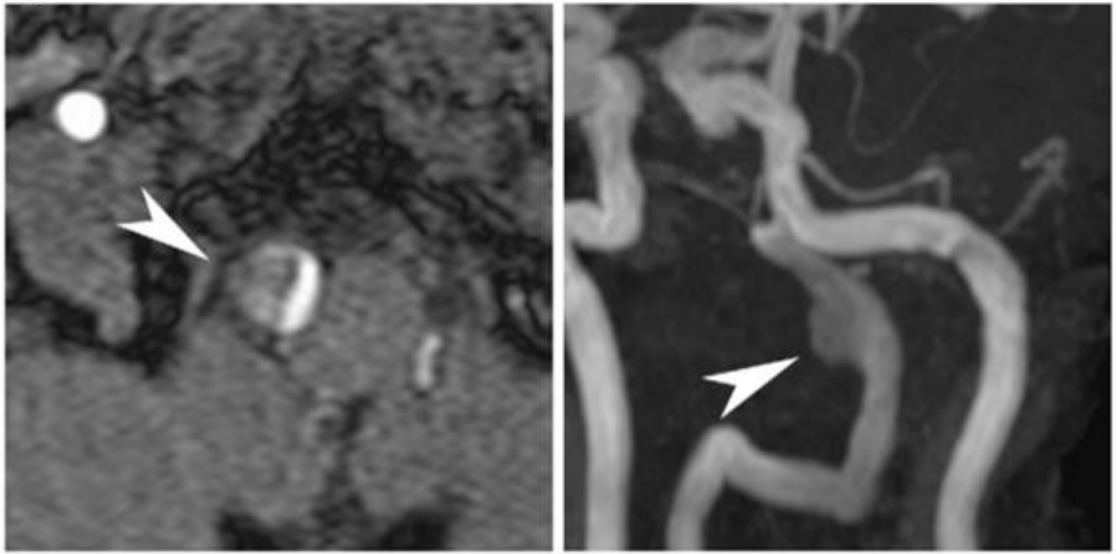
3. **Mini-Batch Gradient Descent**: Compromise btwn BGD & SGD; data divided into small batches and weights/biases updated after processing each batch

Applications of Neural Networks & Other DL Models

1. **Facial Recognition**

Applications of Neural Networks & Other DL Models

2. **Medical Imaging**



Ueda, D., Shimazaki, A. & Miki, Y. Technical and clinical overview of deep learning in radiology. *Jpn J Radiol* **37**, 15–33 (2019). https://doi.org/10.1007/s11604-018-0795-3

Applications of Neural Networks & Other DL Models

3. **Autonomous Driving**
4. **Content Moderation**
5. **Augmented Reality (AR) and Virtual Reality (VR)**

More…

The Math Behind Consciousness?

-Consciousness is required for AGI, according to some scientists.

-Defn: Consciousness is our subjective experience of the brain processing information

The Math Behind Consciousness?

Open Question: How to model consciousness mathematically (non-mathematically)?

-Start From the Origin of Consciousness?

References

-https://link.springer.com/article/10.1007/s11604-018-0795-3#citeas

-https://towardsdatascience.com/the-math-behind-neural-networks-a34a51b93873#03fd

-https://www.ibm.com/topics/deep-learning

-https://www.jasonosajima.com/forwardprop

-https://www.ibm.com/topics/deep-learning

-https://en.wikipedia.org/wiki/Artificial_general_intelligence