ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL & COMPUTATIONAL SCIENCES

Artificial Intelligence for Physics Application

Department of Physics

www.aau.edu..et

mesfin.diro@aau.eud.et

+251912086115 6

EPSNA

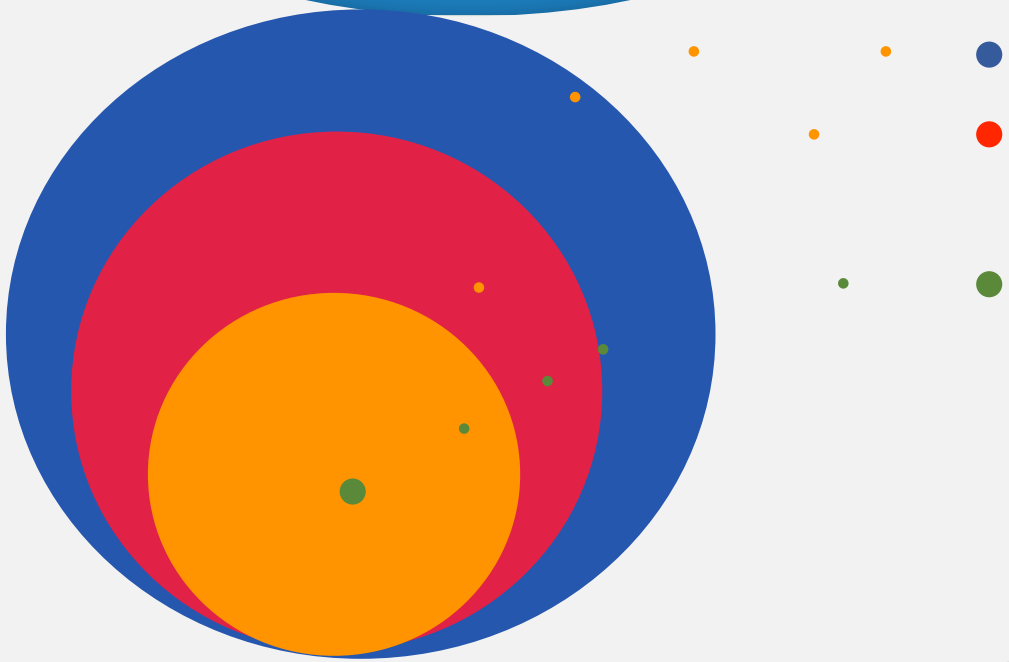
THE ETHIOPIAN PHYSICS SOCIETY IN NORTH AMERICA

Definition of Artificial Intelligence

Artificial Intelligence (AI) means machine or computer which is based upon software algorithm that mimic biologically (naturally) intelligent beings including human cognitive abilities, such as learning, problem-solving, reasoning, planning, natural language processing, and perception

What is AI?

Sub-branches of AI

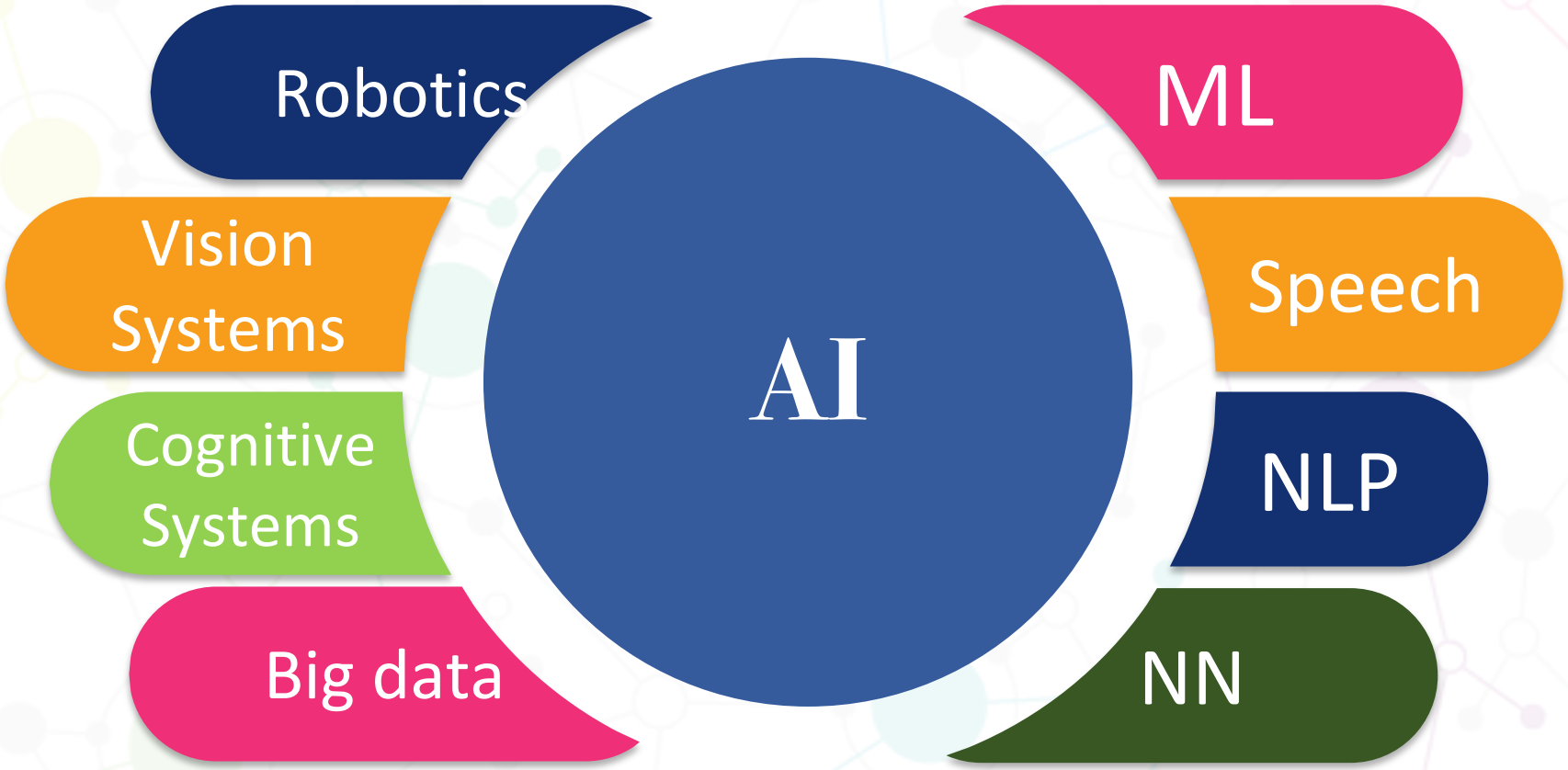


Artificial Intelligence

Machine Learning

Deep Learning

What is AI?


Robotics
ML
Vision Systems
Speech
AI
Cognitive Systems
NLP
Big data
NN

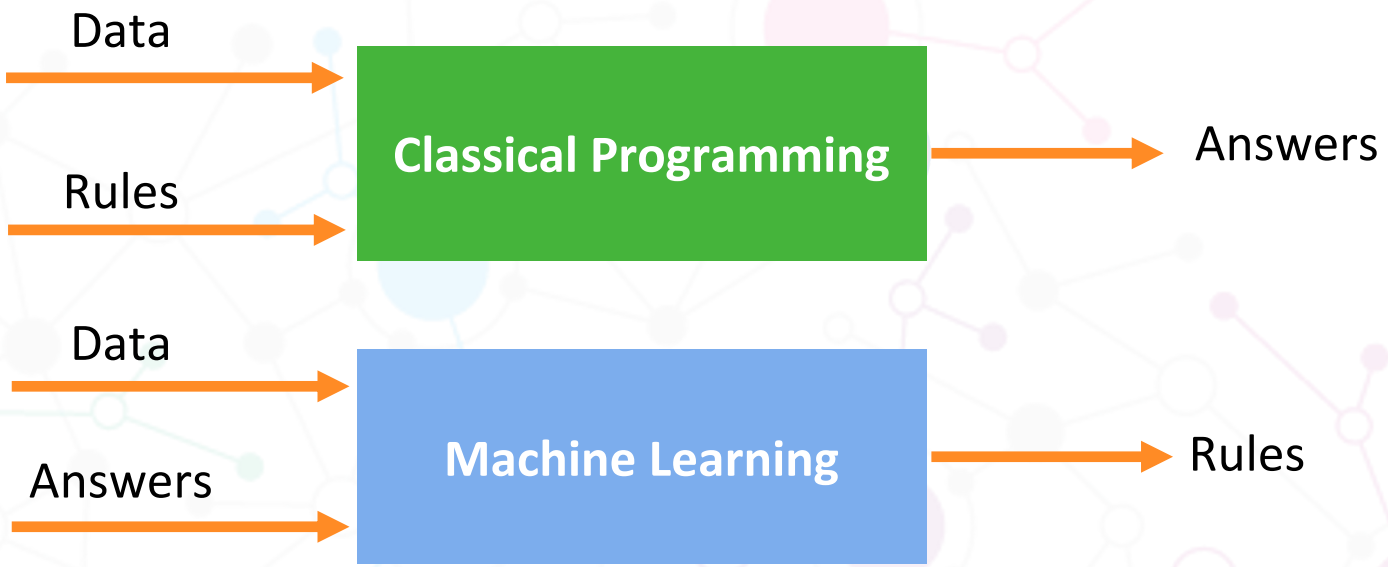
Machine Learning

- Arthor Samuel(1959): Machine Learning is a field of study that gives computers the ability to learn without being explicitly programmed
- Tom Mitchell(1998): A computer program is said to learn from Experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improve with experience E.

ML

Machine Learning

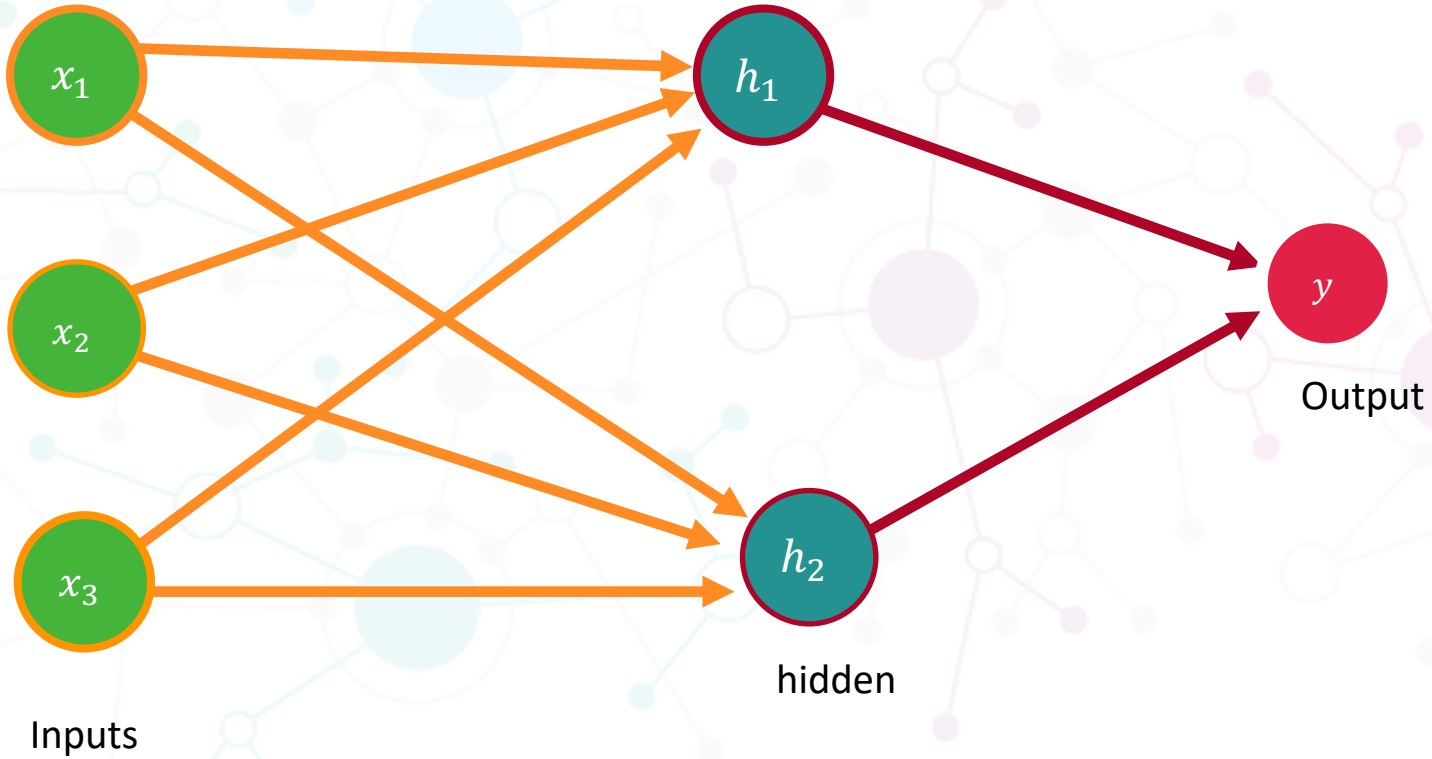
ML



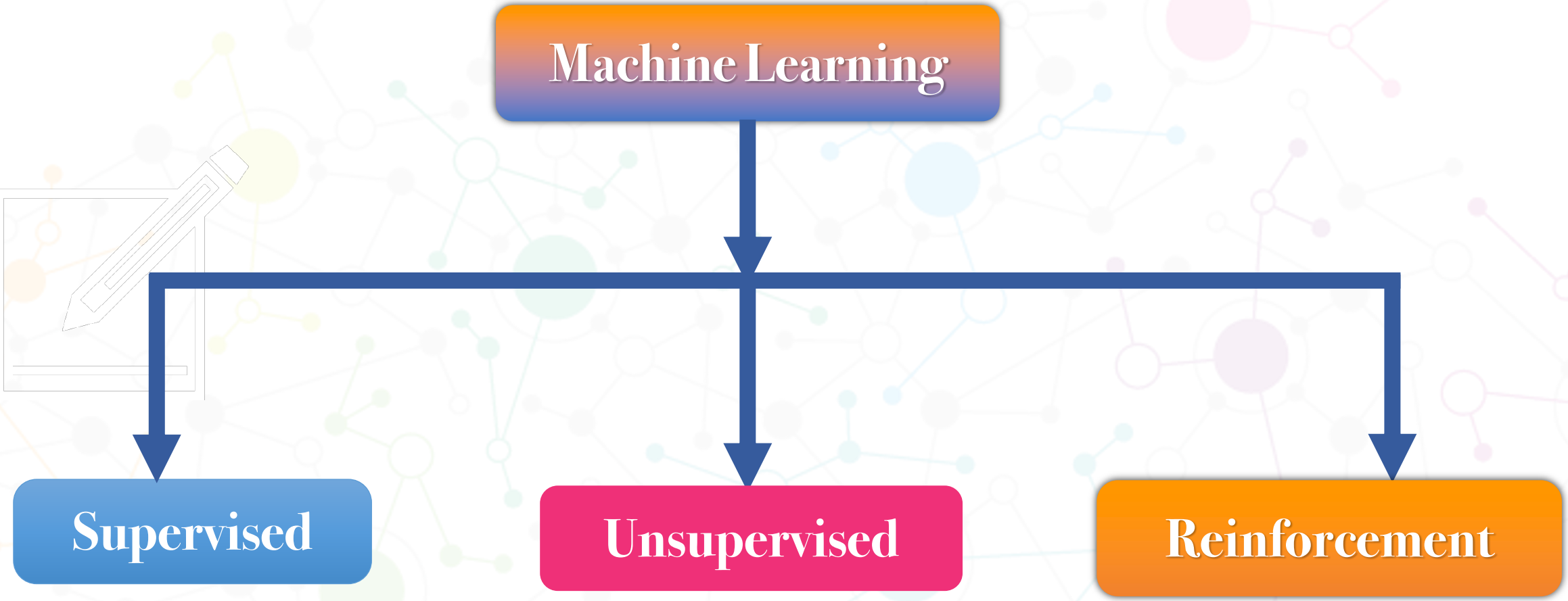
- At a high level, ML systems look a tons of data and come up with rules to predict outcomes for unseen data.
- Fundamentally, machine learning involves building mathematical models to help understand data.

Deep Learning

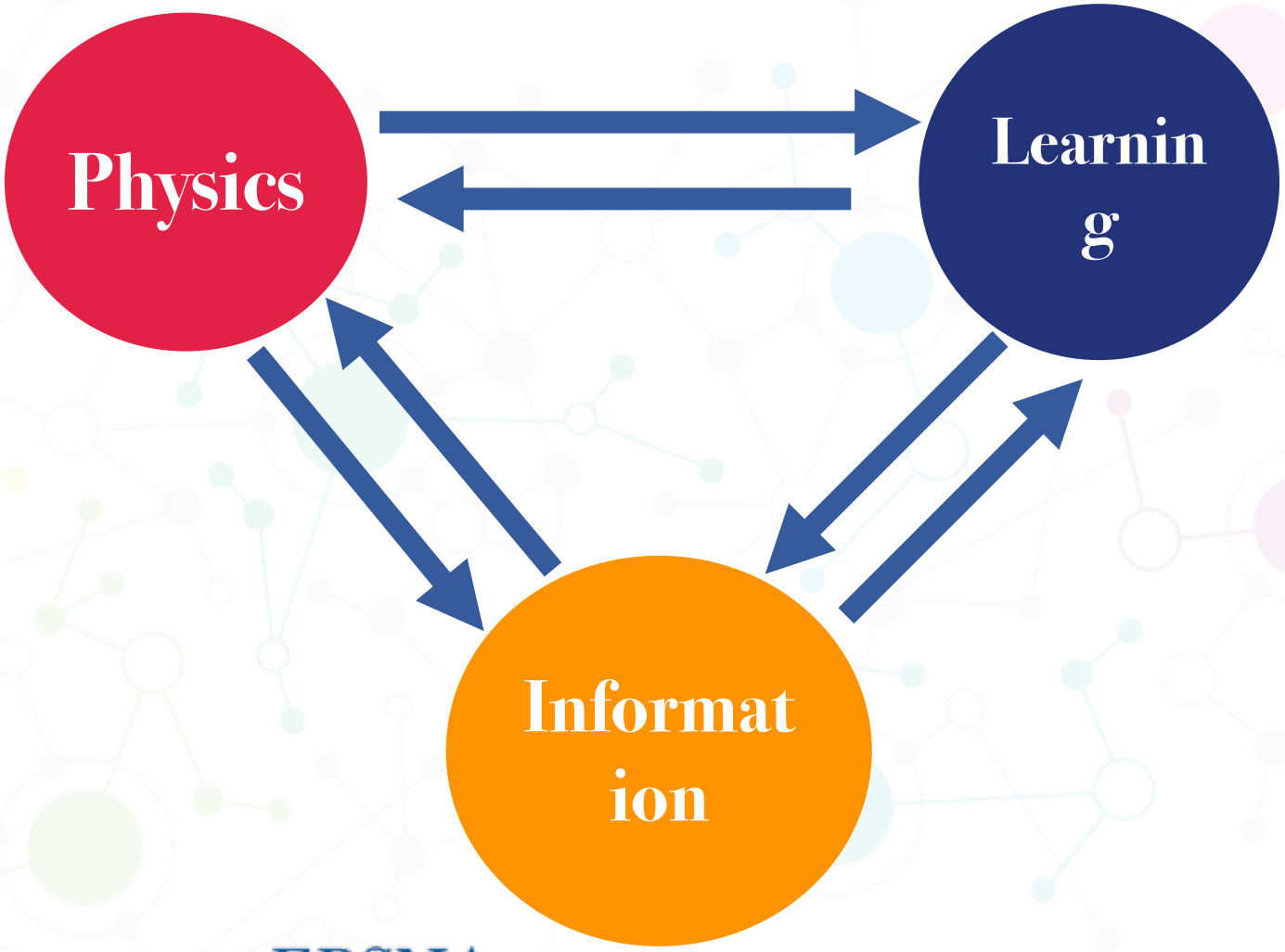
- Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks.



Types of Machine Learning



Physics & Machine Learning



AI for Physics Applications

- Quantum Mechanics
- Statistical Physics
- Material Science
- Nuclear Physics
- Atmospheric Physics
- Particle Physics
- Condensed Matter
- Astrophysics

Quantum Mechanics

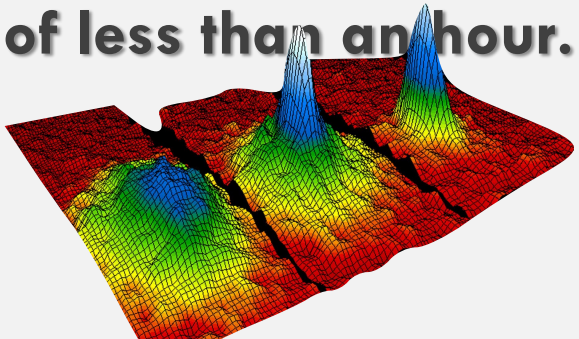
AI

Schrodinger's equation, to find the ground state energy, which is the lowest possible energy of a particle in a one-dimensional box, can be solved with deep learning. Quantum state is a wave-function showing the probability of what is most probable state that the particle can be found in.

Statistical Physics

AI

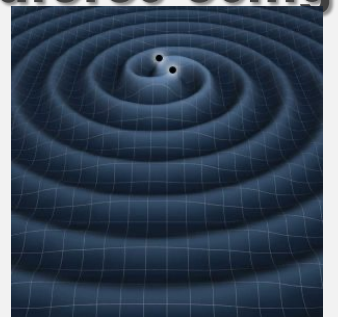
Bose-Einstein condensate, the experiment for which a Nobel Prize was won in 2001, was repeated in 2016, but this time with the aid of AI. The AI could do the whole experiment in a span of less than an hour.



Astrophysics

AI

Gravitational Waves, another Nobel prize winner, is one the most recent biggest discovery in the area of Astrophysics. Studies have been going on find out more Gravitational Wave signatures using AI, via deep learning algorithms.



Atmospheric Physics

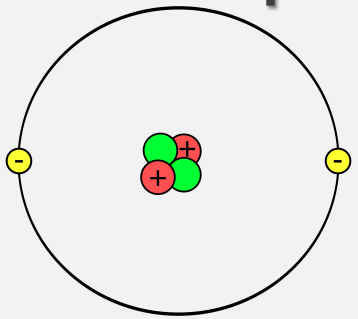
AI

Problems like understanding the mechanism of pollution, identifying cyclones can be addressed using algorithms of AI like Self Organizing Maps and Clustering to cope with these difficult problems.

Condensed Matter Physics

AI

Neural networks are also capable of representing ground state wave-functions. Machine learning tools used in condensed matter and statistical physics to implement various algorithms.



Redox Flow Batteries Application

- Renewable-energy sources, such as solar and wind, are being deployed in larger numbers than ever before(Weber et al., 2011)
- Wind and Solar energy sources are regarded as the promising alternatives to the future energy resources due to their environmental friendly characteristics.
- To ensure the resilience of intermittent nature of these sources , advanced electrical energy storage devices needed for the electrical grid system to level out the irregularities in power supply (Cheng et al., 2015)

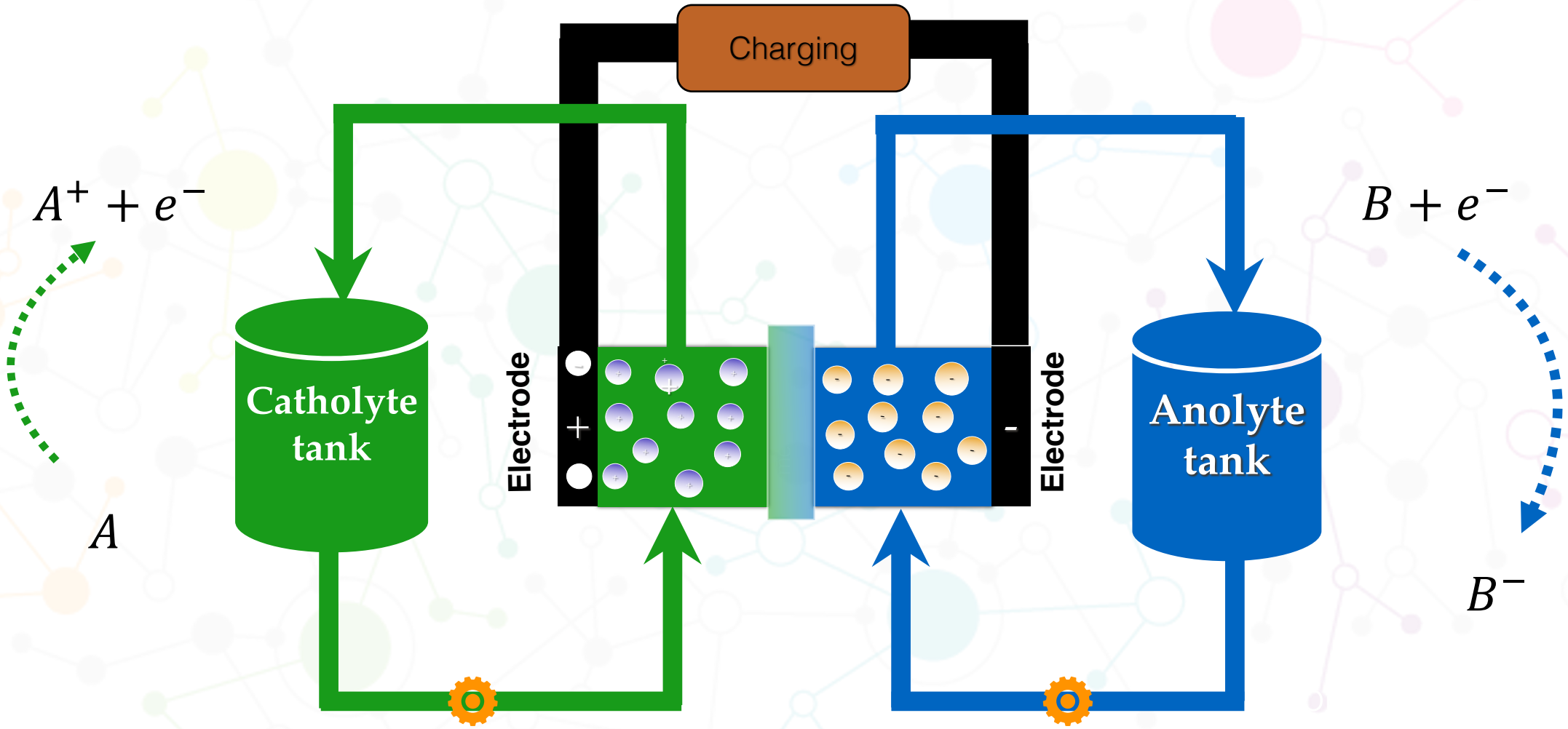
Redox Flow Batteries Application

- Redox flow batteries (RFBs) have emerged as prime candidates for energy storage on the medium and large scales, particularly at the grid-scale due to the demand for versatile energy storage for electrical energy generated from intermittent renewable sources.
- RFBs substance react in a reversible manor on electrodes though energy conversion in an electrochemical cells.

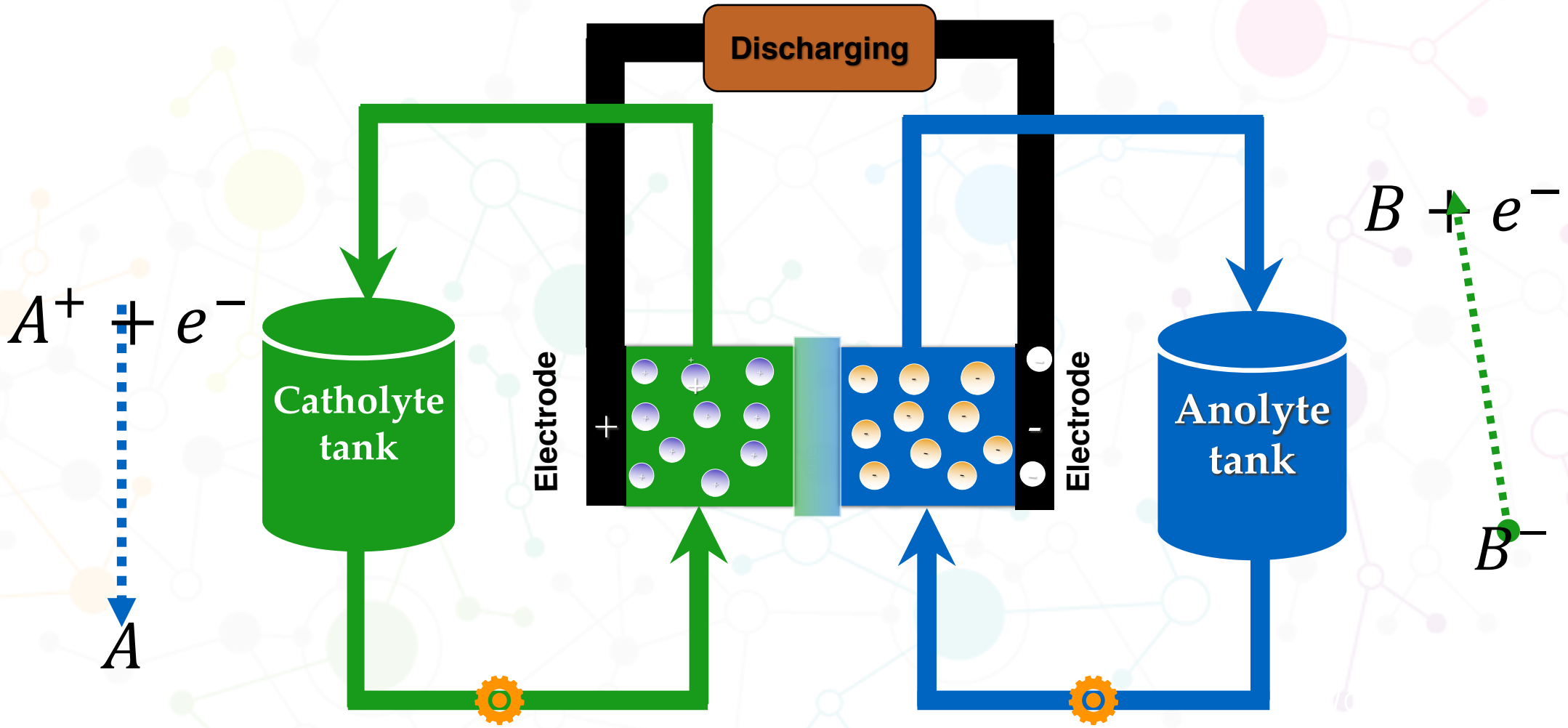
Redox Flow Batteries Application

- The development of RFBs based on inexpensive and sustainable organic redox-active materials can overcame the drawbacks of metal-based flow batteries
- However, the chemical search space of organic molecules is extremely large in the order of 10^{100} (drug-like, Photovoltaic, Polymers & dyes) . So far, only 10^{8} chemical space are concurred based on the researchers intuitions and previous research outputs.
- These molecules are also discrete in nature which prevents the use of gradient-based optimization order to accelerate the traditional computational methods.
- The aforementioned opportunities and challenges, the recent advancement of machine learning algorithms and availability of large organic molecules are some the motivation for this study.

Materials Discovery with Deep Learning

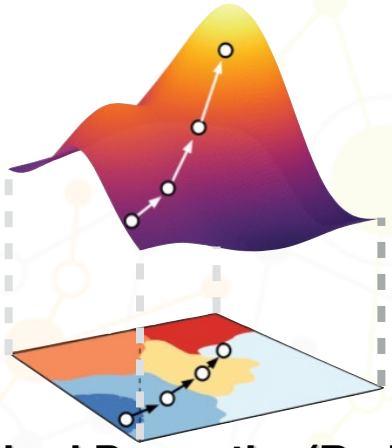


Materials Discovery with Deep Learning



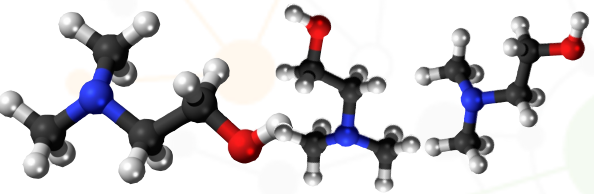
Materials Discovery with Deep Learning

Functional Space



Desired Properties(Redox Potential,Solubility, Stability)

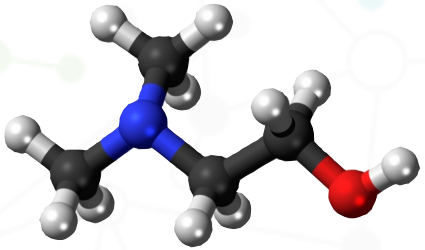
Chemical Space



drug-like, Photovoltaic, Polymers & dyes

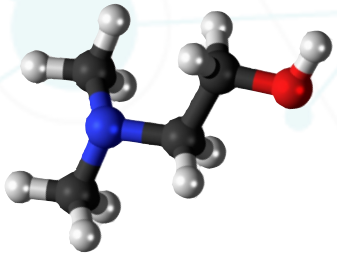
Direct

Expt or Simulation

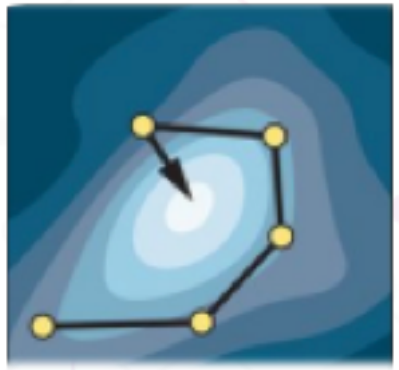


Inverse

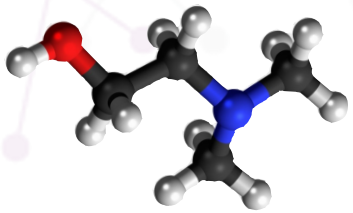
High-throughput



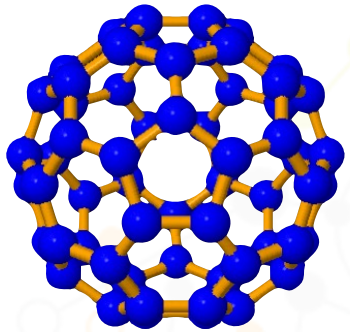
Inverse



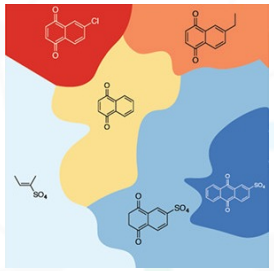
GAN, VAE etc



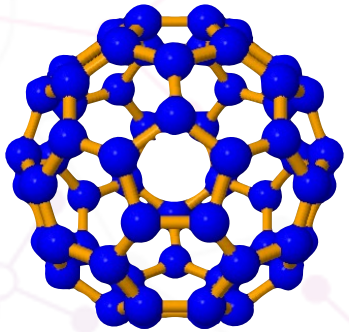
Materials Discovery with Deep Learning



Encoder



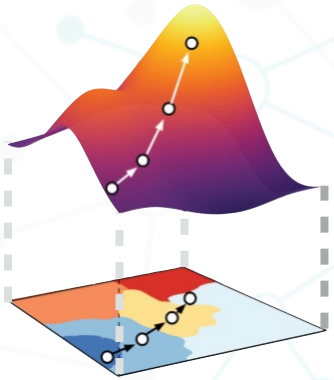
Decoder



Latent Space

Molecular Representations (SMILES, Graph, 3D)

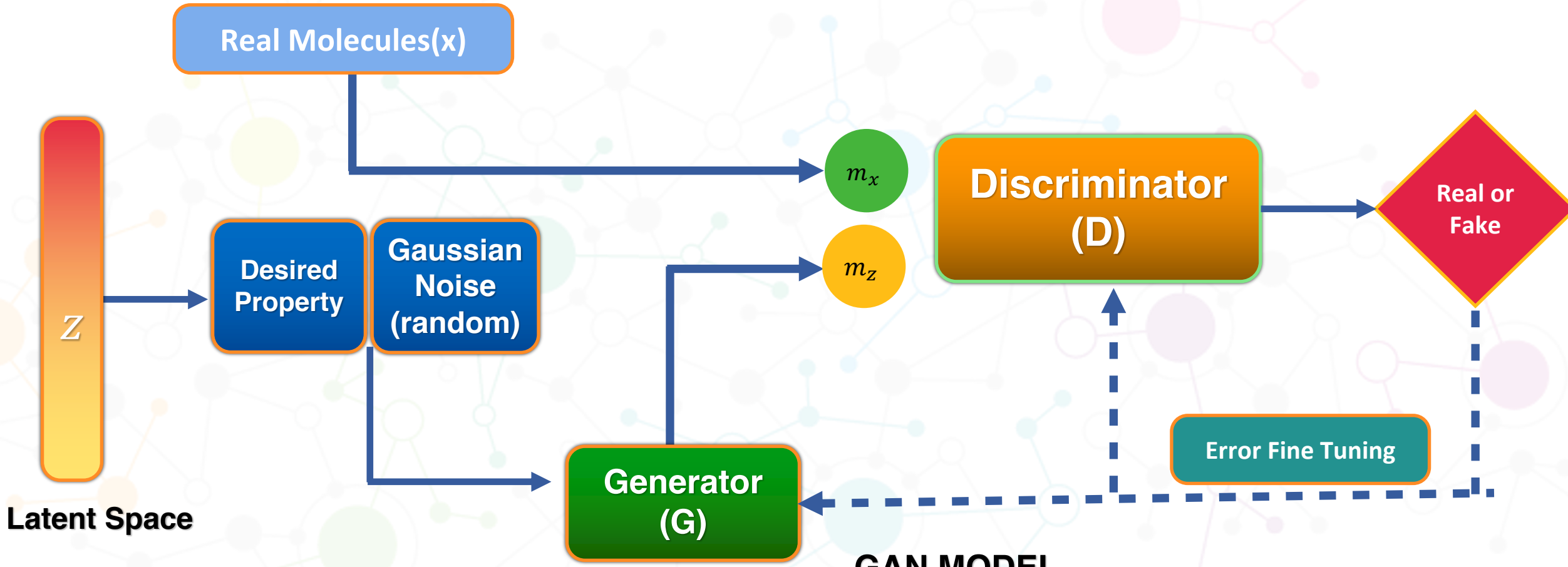
Generated Molecule



VAE MODEL REPRESENTATION

Property Prediction (Redox Potential, Solubility, Stability)

Materials Discovery with Deep Learning



GAN MODEL REPRESENTATION

Loss Function in GAN

- Let's define some notations that will be applied to our generative adversarial network model:
  - E_x: is the expected value over all real molecular data instances
  - E_z: is the expected value over all random inputs to the generator
  - x : Real molecular data , z: Latent vector
  - E_z: is the discriminator's estimate of the probability that a fake molecular instance is real
  - $D(x)$ is the discriminator's estimate of the probability that real molecular data insta nce x
  - $D(G(z))$: Discriminator evaluation on the generator's output is real

Loss Function in GAN

- The objective of the discriminator is to correctly label generated molecular data as false and empirical data points as real. Therefore, the loss function of the discriminator is:

$$\max_{D} V(G,D) = loss(D(x),1) + loss(D(G(z)),0)$$

$$= -E_{x\sim p(x)}[\log D(x)] - E_{z\sim p(z)}[\log(1 - D(G(x)))]$$

Objective Function

- Similarly the objective of the generator is to confuse the discriminator as much as possible such that it mislabels generated molecules as being true with a loss after applying binary-entropy:

$$\min_{G} V(G,D) = loss(D(G(z)),1)$$

$$= E_{z\sim p(z)}[-\log(D(G(z)))]$$

Loss Function in GAN

- **In general the discriminator model $D(x)$ and the generator model $G(z)$ are playing minmax zero-sum game with the comprehensive objective function from cross-entropy distribution is:**

$$\min_{G}\max_{D} V(G,D) = E_x[\log D(x)] + E_z[\log(1 - D(G(z))]$$

- D is accurate over real data by maximizing $E_x[\log D(x)]$
- G is trained to increase D's probability for fake molecules by minimizing $E_z[\log(1 - D(G(z))]$

Molecular Data Sources

- Deep generative models require large datasets in order to learn patterns that can generalize and lead to new molecules:
  - GDB 17 dataset - 166 Billion (Ruddigkeit et al., 2012)
  - GDB 13 dataset - 970 Million (Blum and Reymond, 2009)
  - QM9 Dataset - 134 kilo (Ramakrishnan et al., 2014)
  - ChemBL database (Gaulton et al., 2017)
  - ZINC database - 980 Million (Irwin and Shoichet, 2005)
  - MolData Dataset

Molecular Data Representation

Fingerprints

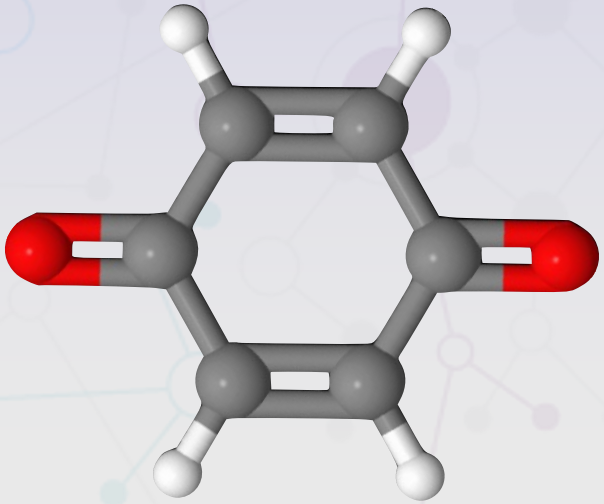
0000000100100000010001000000
0000000110010000000000000000
000100000

SMILIES

O=C1C=CC(=O)C=C1

Graph

3D



T H A N K YOU!

EPSNA

THE ETHIOPIAN PHYSICS SOCIETY IN NORTH AMERICA