UE21CS343BB2

Topics in Deep Learning

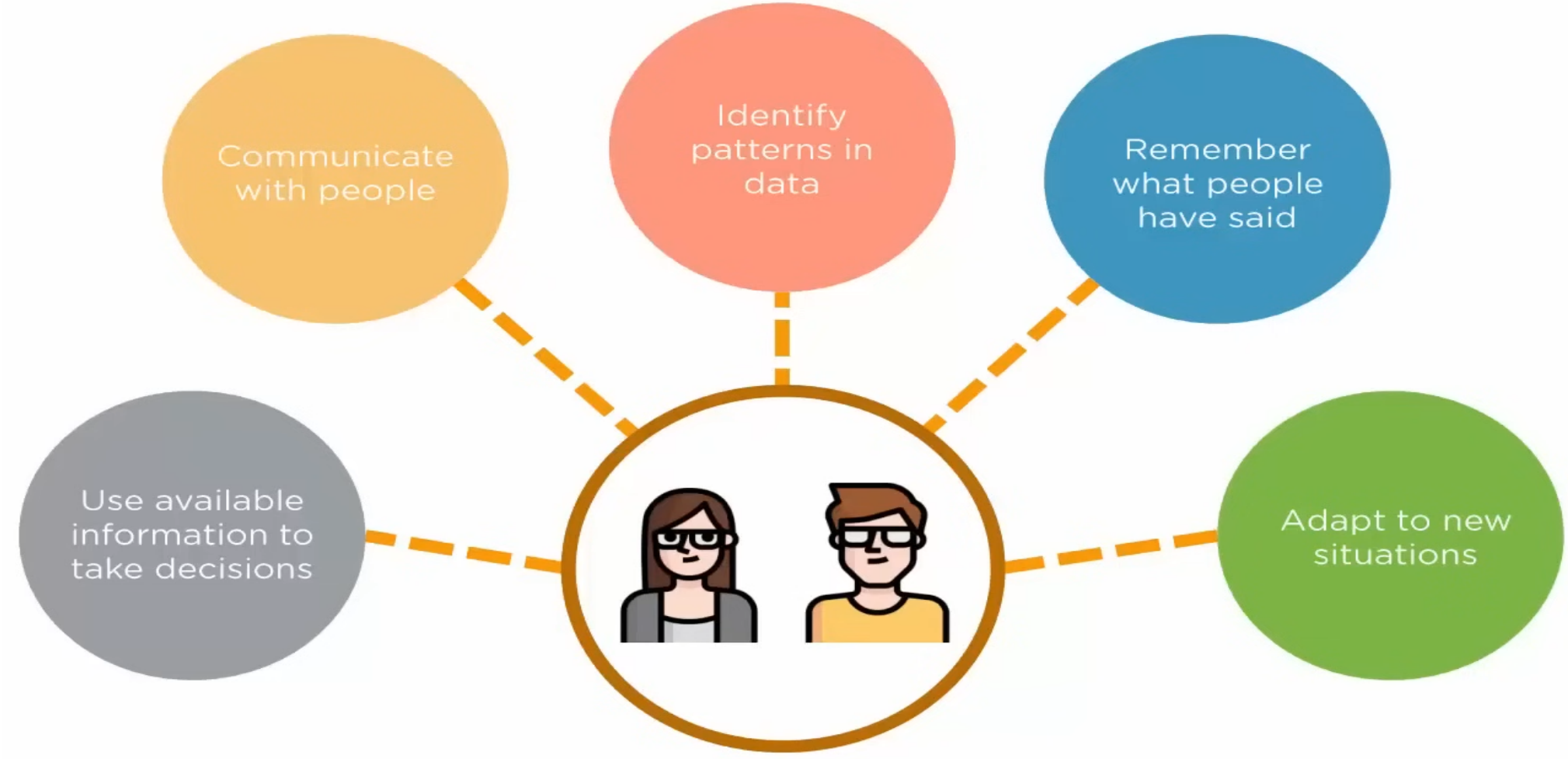
Dr. Shylaja S S
Director of Cloud Computing & Big Data (CCBD),
Centre for Data Sciences & Applied Machine Learning (CSSAML)
Department of Computer Science and Engineering
shylaja.sharath@pes.edu

PES UNIVERSITY

Topics in Deep Learning

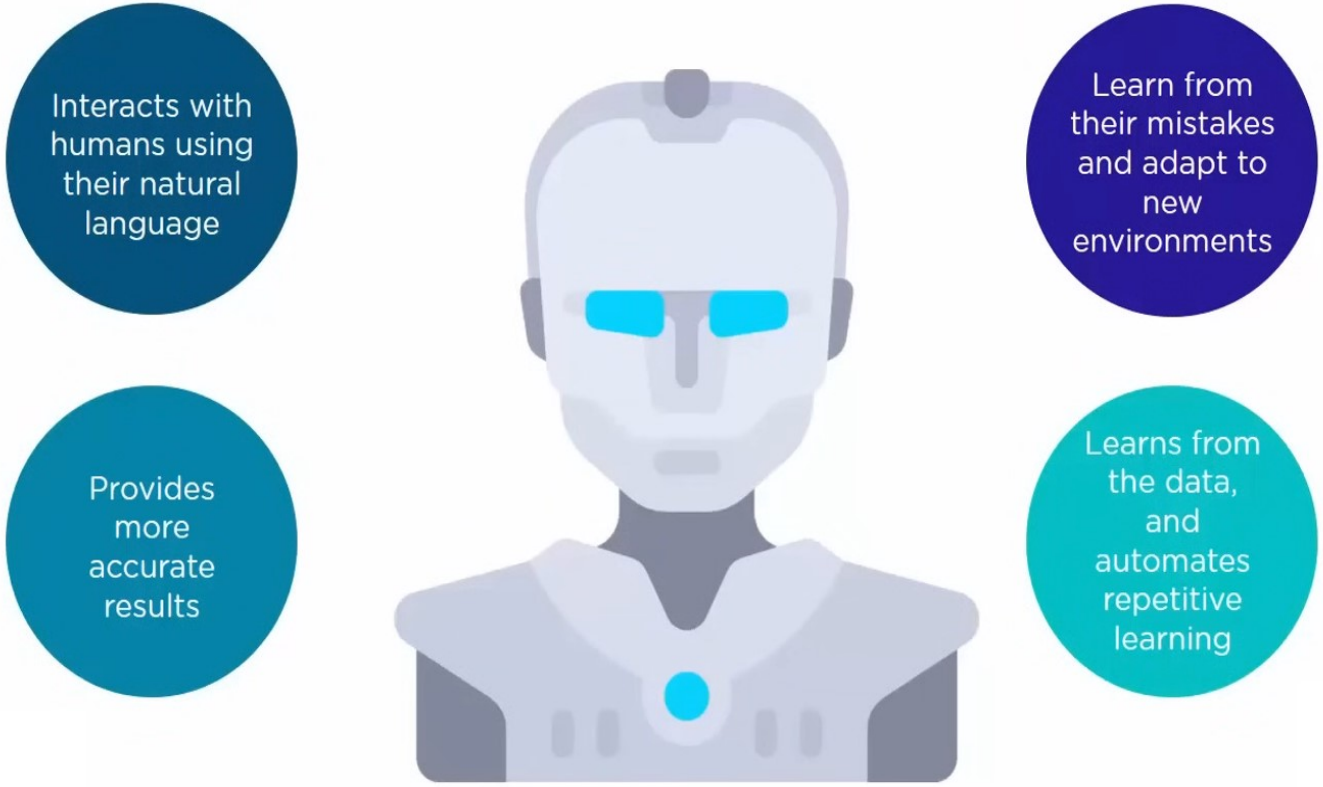
Introduction to Deep Learning

Human Intelligence

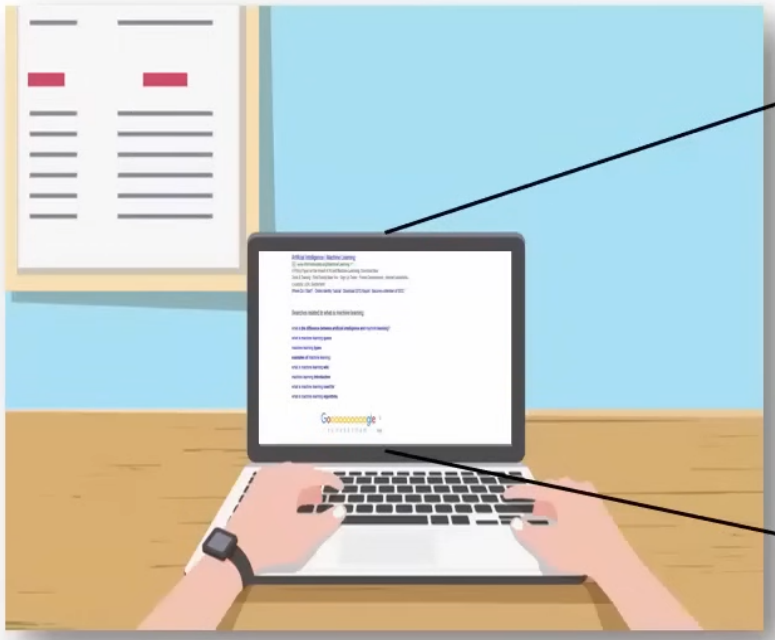


Artificial Intelligence

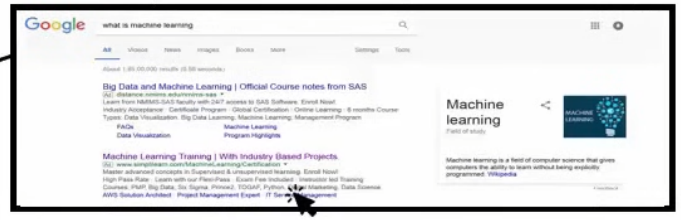
AI develops computer systems that can accomplish tasks that require human intelligence



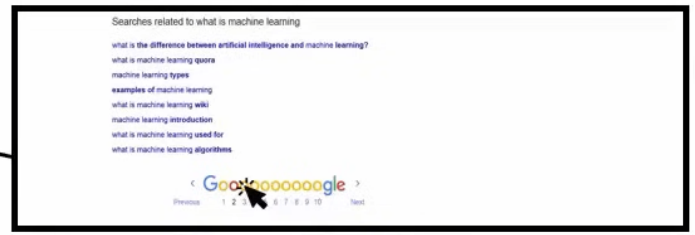
Machine Learning



User searches for something on Google



User selects one of the first few links and spends time there



User goes to the second/ third page

Google understands the user got what was required

Google understands the user's requirement wasn't satisfied

Types of Machine Learning

Supervised Learning

- Makes machine Learn explicitly
- Data with clearly defined output is given
- Direct feedback is given
- Predicts outcome/future
- Resolves classification and regression problems

Unsupervised Learning

- Machine understands the data (Identifies patterns/structures)
- Evaluation is qualitative or indirect
- Does not predict/find anything specific

Reinforcement Learning

- An approach to AI
- Reward based learning
- Learning form +ve & +ve reinforcement
- Machine Learns how to act in a certain environment
- To maximize rewards

Types of Machine Learning

ARTIFICIAL INTELLIGENCE

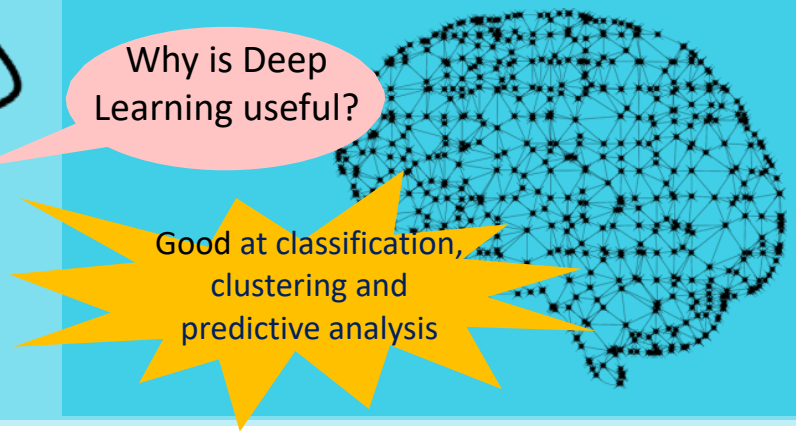
Broad area which enables computers to mimic human behavior

MACHINE LEARNING

Usage of statistical tools enables machines to learn from experience (data) – need to be told

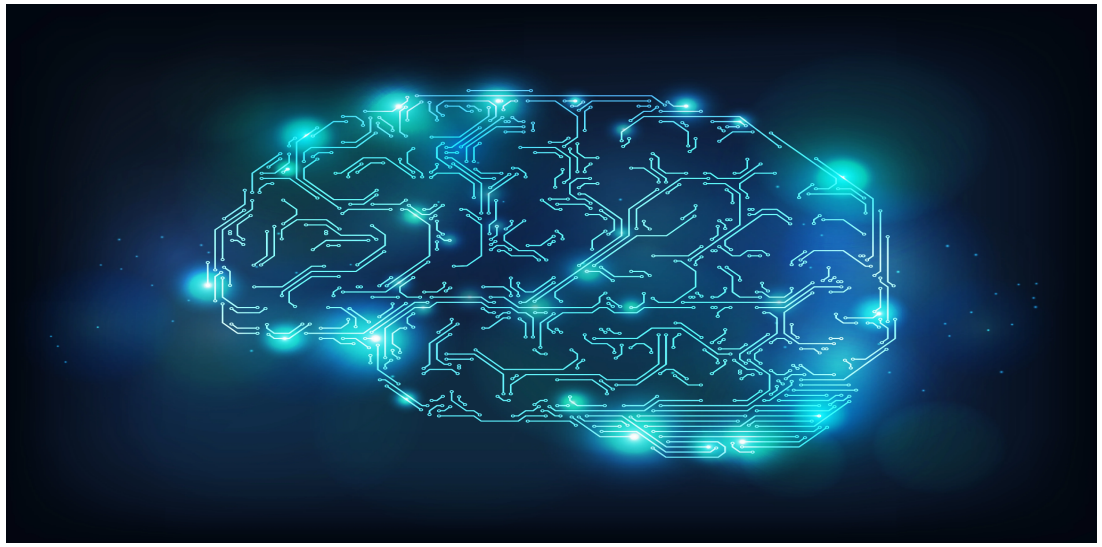
DEEP LEARNING

Learn from its own method of computing - its own brain



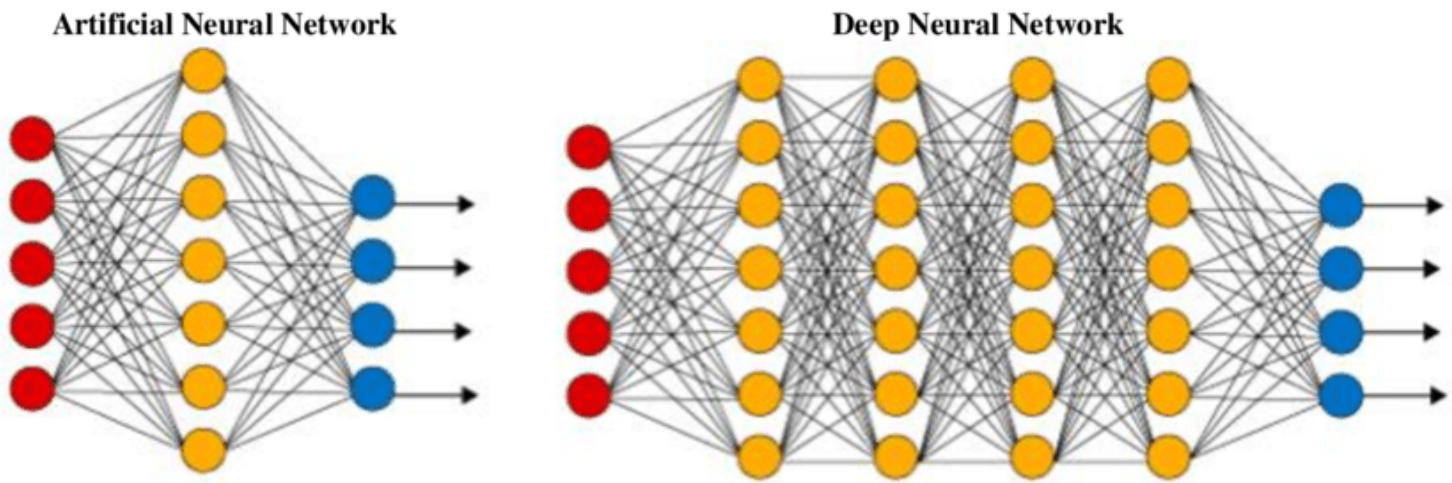
Deep Learning

- Deep learning is a branch of machine learning which is based on artificial neural networks.
- It is capable of learning complex patterns and relationships within data and does not require us to explicitly program everything.
- It has become increasingly popular in recent years due to the advances in processing power and the availability of large datasets.



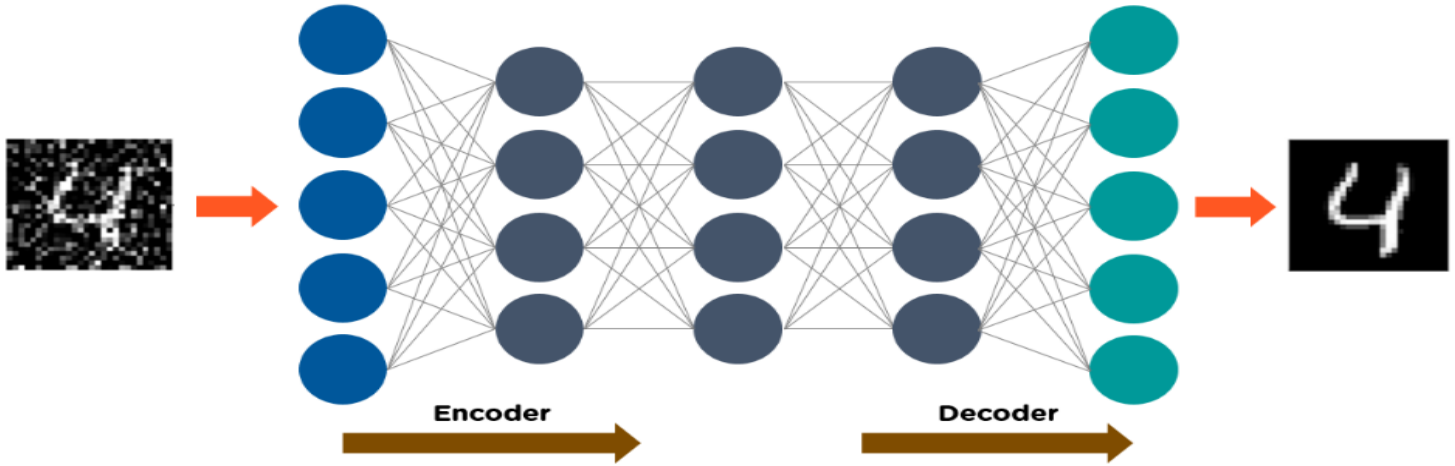
Deep Learning

- An artificial neural network or ANN uses layers of interconnected nodes called neurons that work together to process and learn from the input data.
- In a fully connected Deep neural network, there is an input layer and one or more hidden layers connected one after the other.
- Each neuron receives input from the previous layer neurons or the input layer.



Deep Learning

- The output of one neuron becomes the input to other neurons in the next layer of the network, and this process continues until the final layer produces the output of the network.
- The layers of the neural network transform the input data through a series of nonlinear transformations, allowing the network to learn complex representations of the input data.



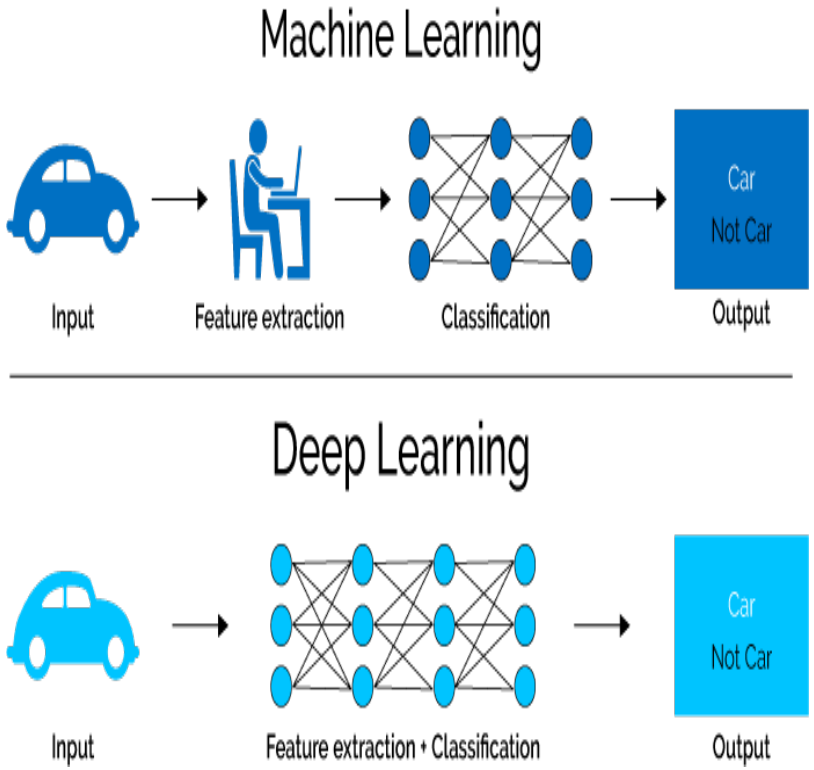
Machine Learning vs Deep Learning

- **Automatic Feature Extraction:** Deep learning models, especially neural networks, automatically learn relevant features from raw input data during the training process.
- **Elimination of Manual Feature Engineering:** Unlike traditional machine learning, deep learning reduces the need for manual feature extraction, saving time and leveraging the model's ability to learn complex patterns.
- **Hierarchical Representations:** Deep neural networks consist of multiple layers that transform input data into hierarchical and abstract representations, capturing intricate features without explicit human guidance.

Machine Learning vs Deep Learning

- **Applicability to Complex Data:** Deep learning excels in tasks like image recognition, natural language processing, and speech recognition, where the data's complexity may be challenging to capture through manual feature engineering.

While deep learning does automate feature extraction, attention to proper data preprocessing and hyperparameter tuning is crucial for achieving optimal model performance.



Machine Learning vs Deep Learning

Machine Learning	Deep Learning
Apply statistical algorithms to learn the hidden patterns and relationships in the dataset.	Uses artificial neural network architecture to learn the hidden patterns and relationships in the dataset.
Can work on the smaller amount of dataset.	Requires the larger volume of dataset compared to machine learning.
Better for the low-label task.	Better for complex task like image processing, natural language processing, etc.
Takes less time to train the model.	Takes more time to train the model.
A model is created by relevant features which are manually extracted from images to detect an object in the image.	Relevant features are automatically extracted from images. It is an end-to-end learning process.
Less complex and easy to interpret the result.	More complex, it works like the black box interpretations of the result are not easy.
It can work on the CPU or requires less computing power as compared to deep learning.	It requires a high-performance computer with GPU.

Applications of Deep Learning

- **Computer Vision**

Deep learning models can enable machines to identify and understand visual data. Some of the main applications of deep learning in computer vision include:

- **Object detection and recognition:** Deep learning model can be used to identify and locate objects within images and videos, making it possible for machines to perform tasks such as self-driving cars, surveillance, and robotics.
- **Image classification:** Deep learning models can be used to classify images into categories such as animals, plants, and buildings. This is used in applications such as medical imaging, quality control, and image retrieval.
- **Image segmentation:** Deep learning models can be used for image segmentation into different regions, making it possible to identify specific features within images.

Applications of Deep Learning

- **Computer Vision**

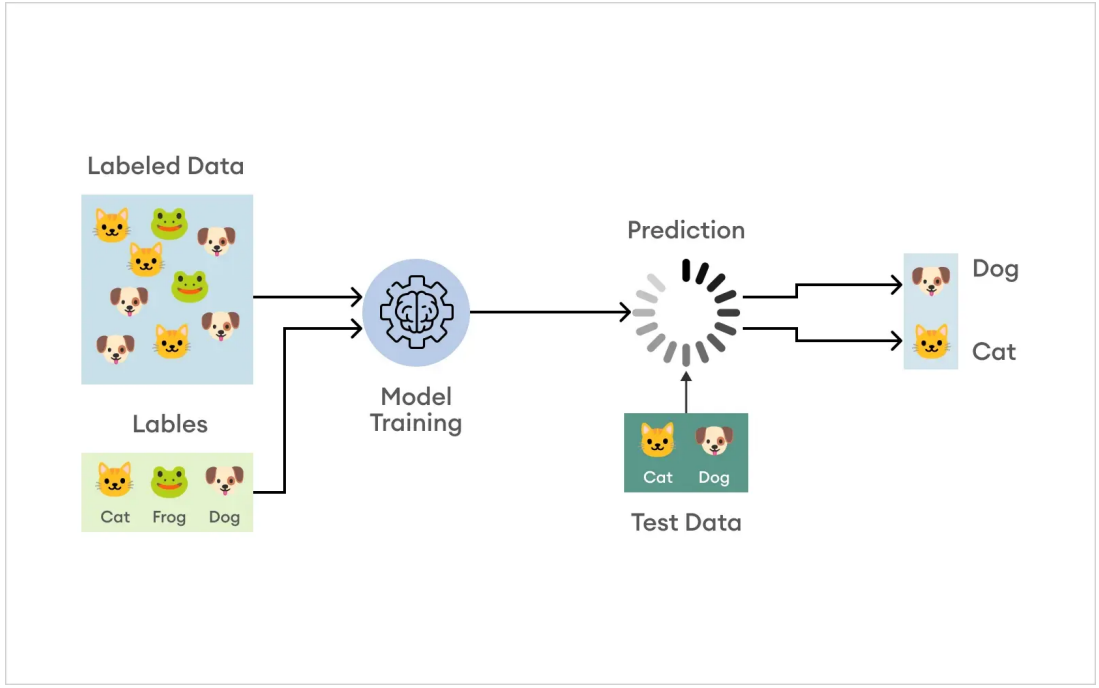


Image Classification

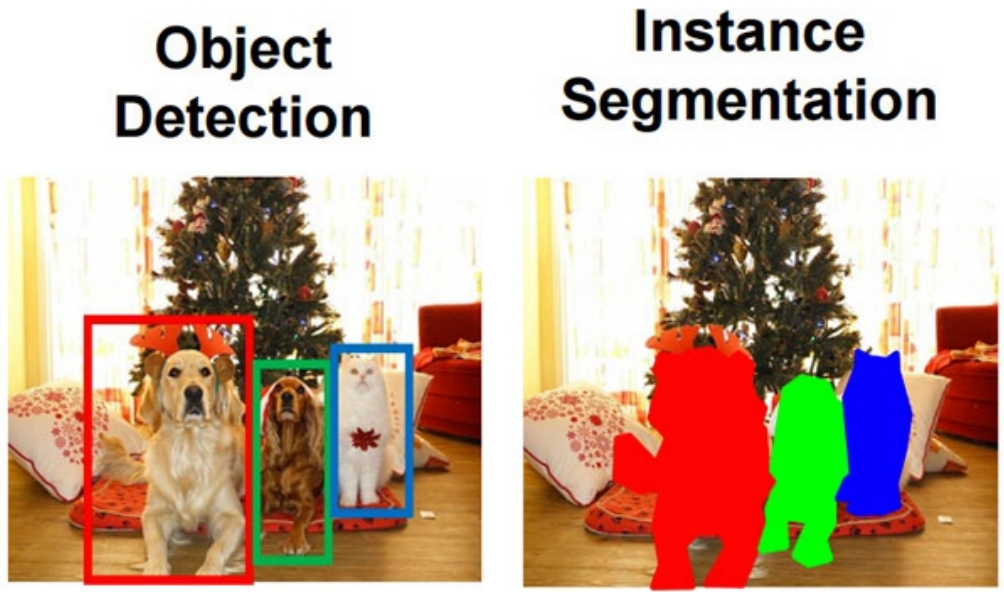


Image Segmentation

Applications of Deep Learning

- **Natural Language Processing(NLP)**

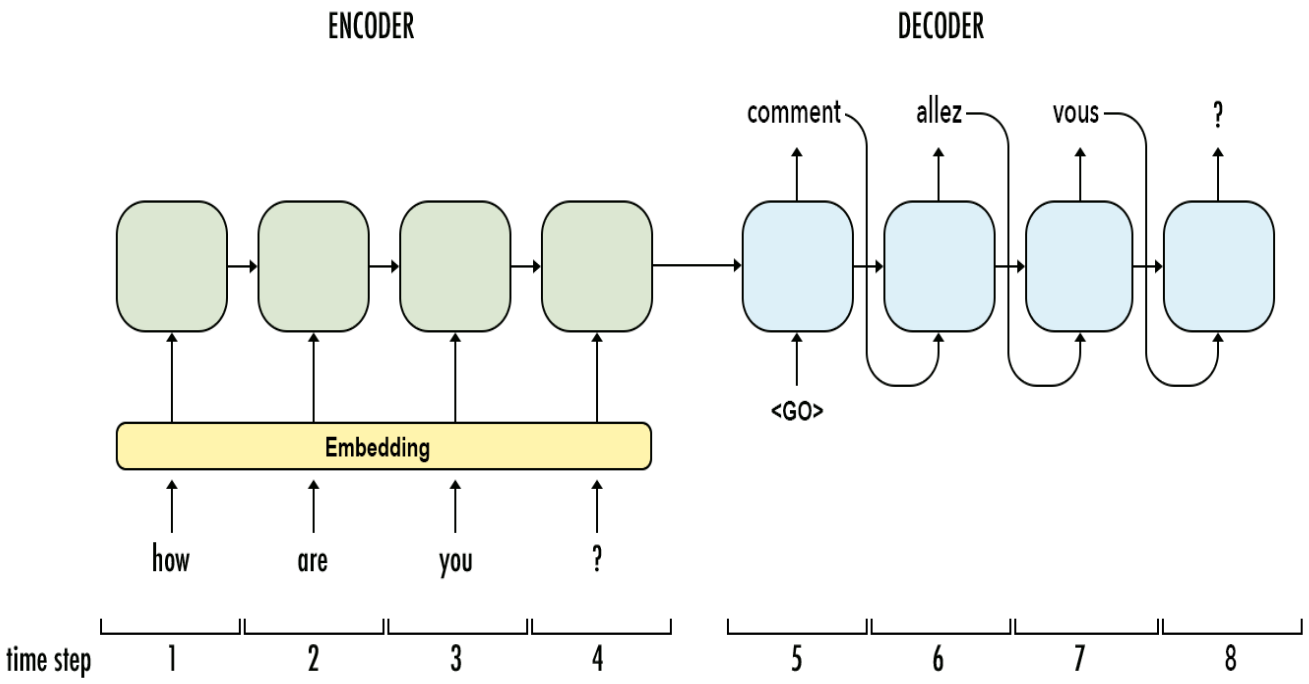
In NLP, the Deep learning model can enable machines to understand and generate human language.

Some of the main applications of deep learning in NLP include:

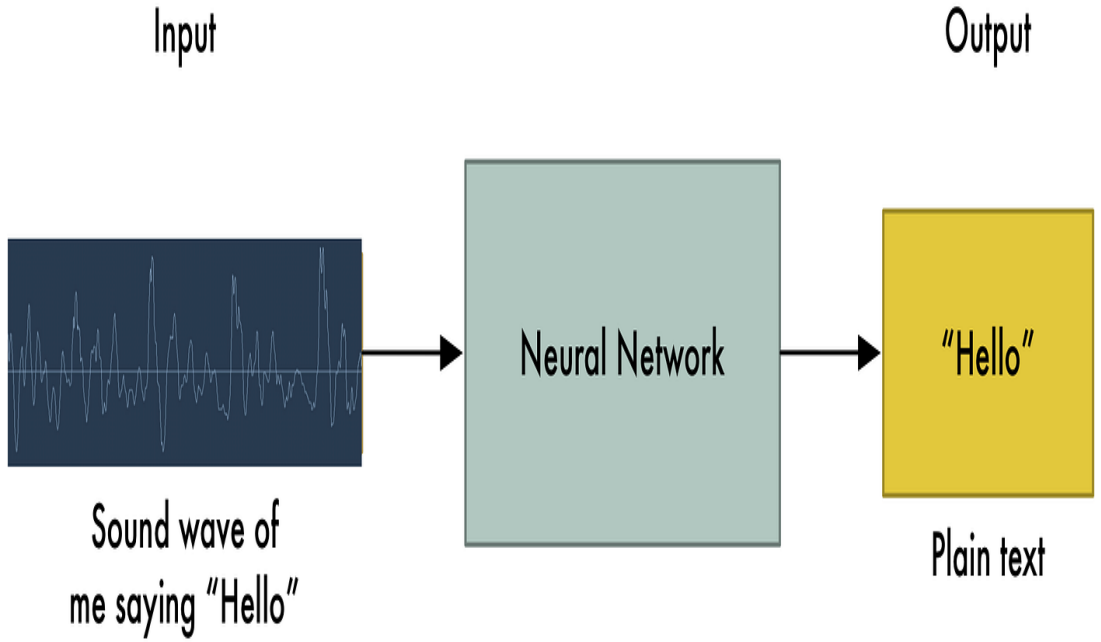
- **Automatic Text Generation**: Deep learning model can learn the corpus of text and new text like summaries, essays can be automatically generated using these trained models.
- **Language translation:** Deep learning models can translate text from one language to another, making it possible to communicate with people from different linguistic backgrounds.
- **Sentiment analysis:** Deep learning models can analyze the sentiment of a piece of text, making it possible to determine whether the text is positive, negative, or neutral. This is used in applications such as customer service, social media monitoring, and political analysis.
- **Speech recognition:** Deep learning models can recognize and transcribe spoken words, making it possible to perform tasks such as speech-to-text conversion, voice search, and voice-controlled devices.

Applications of Deep Learning

- **Natural Language Processing(NLP)**



Language Translation



Speech Recognition

Applications of Deep Learning

- **Reinforcement Learning**

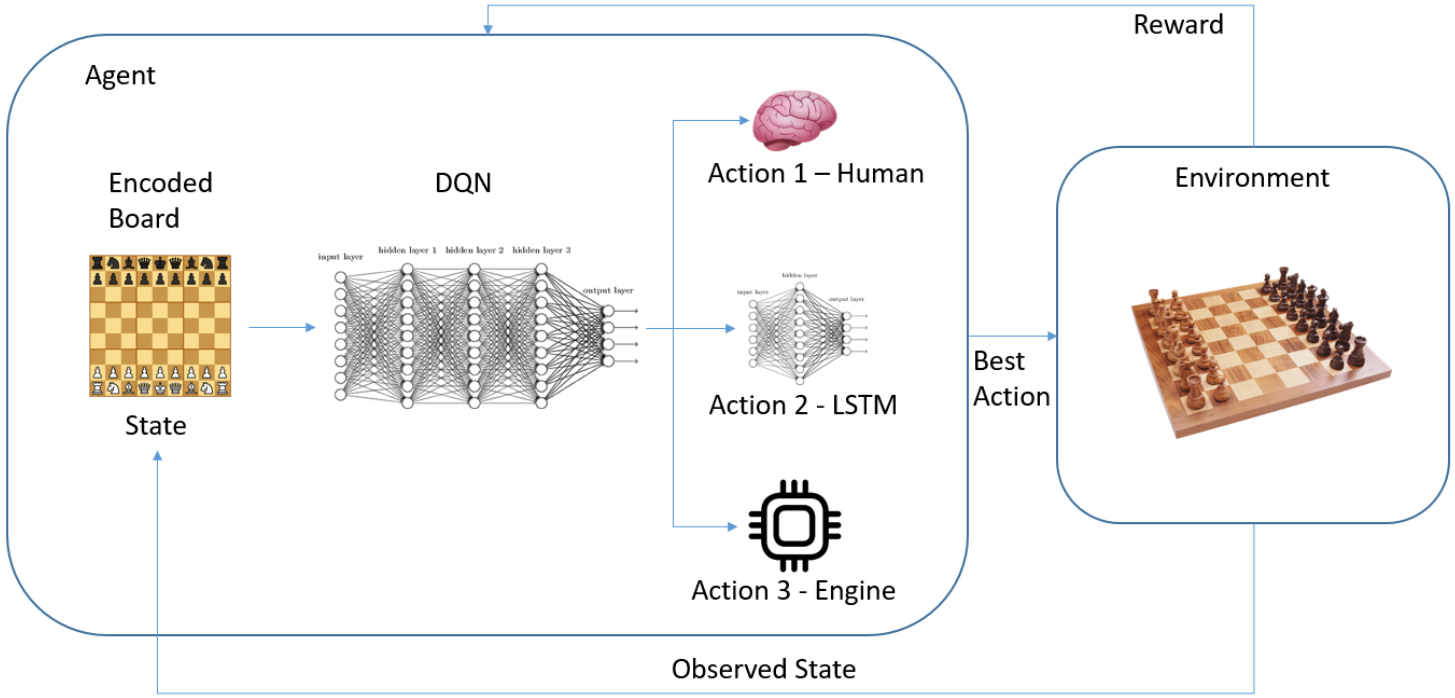
In reinforcement learning, deep learning works as training agents to take action in an environment to maximize a reward.

Some of the main applications of deep learning in reinforcement learning include:

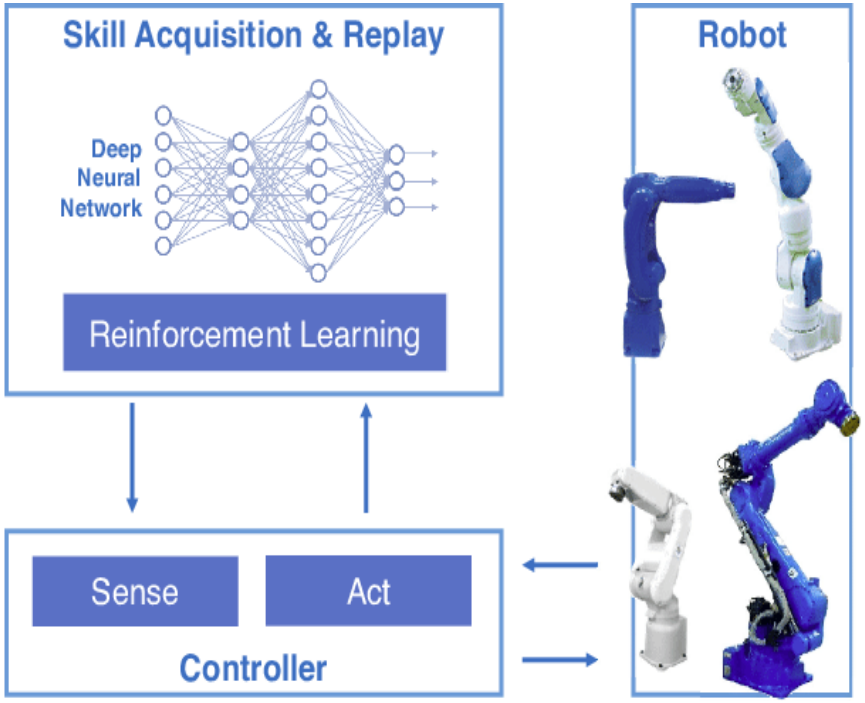
- **Game playing:** Deep reinforcement learning models have been able to beat human experts at games such as Go, Chess, and Atari.
- **Robotics:** Deep reinforcement learning models can be used to train robots to perform complex tasks such as grasping objects, navigation, and manipulation.
- **Control systems:** Deep reinforcement learning models can be used to control complex systems such as power grids, traffic management, and supply chain optimization.

Applications of Deep Learning

- **Reinforcement Learning**



Game Playing



Robotics

Popular Architectures

- **Convolutional Neural Networks**

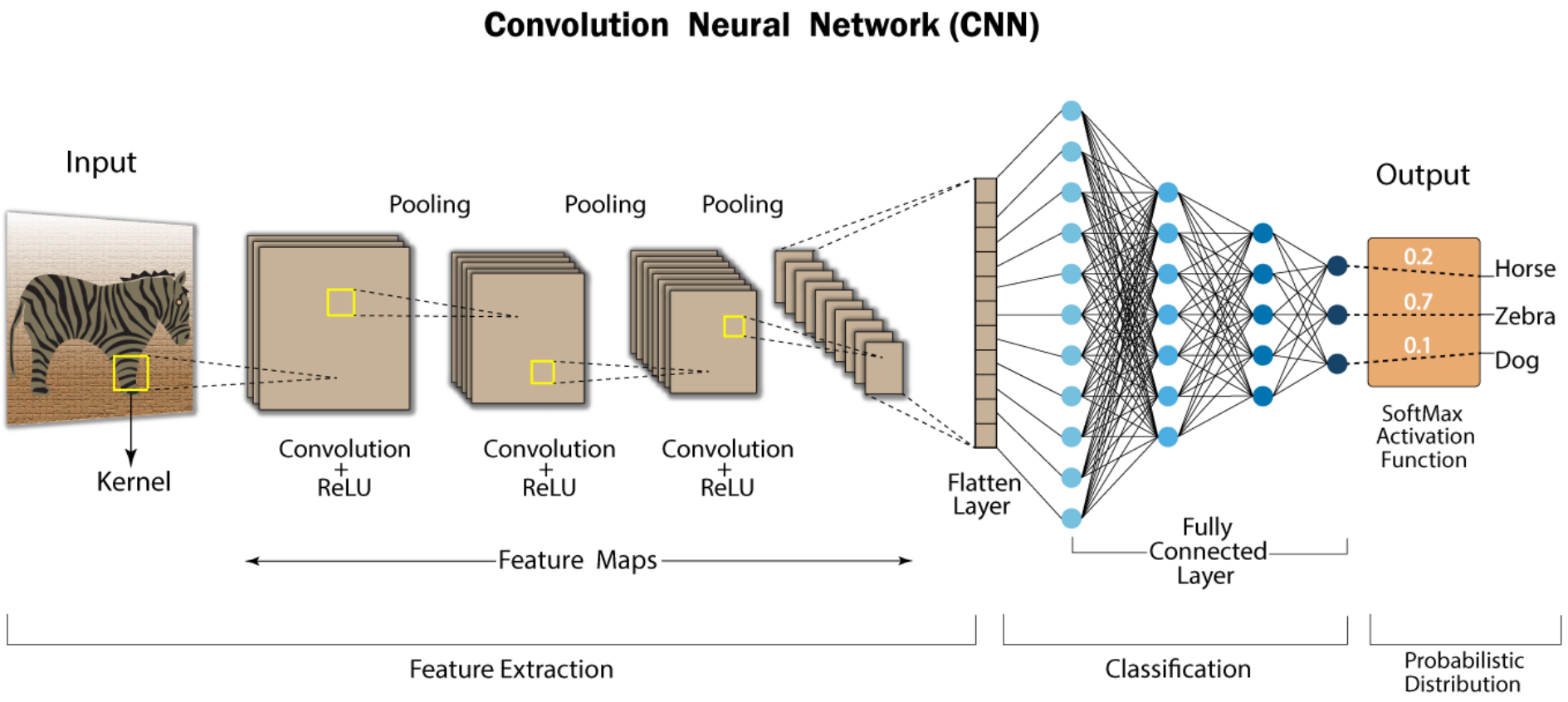
CNNs are a specialized type of neural network designed for processing structured grid data, such as images.

A CNN captures the spatial features from an image. Spatial features refer to the arrangement of pixels and the relationship between them in an image. They help us in identifying the object accurately, the location of an object, as well as its relation with other objects in an image.

For example, face and object recognition software.

Popular Architectures

- **Convolutional Neural Networks**



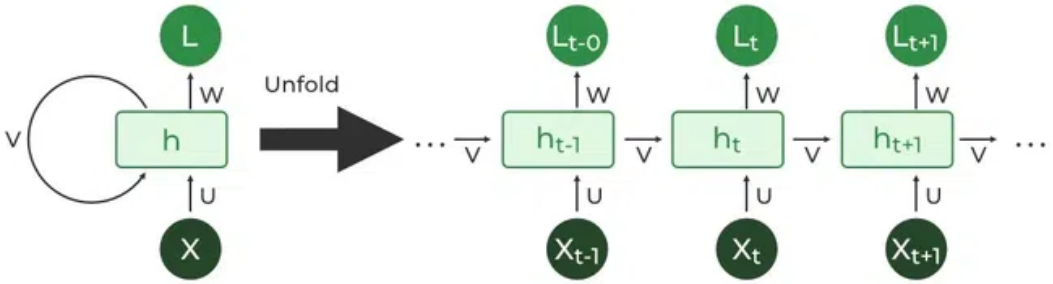
Popular Architectures

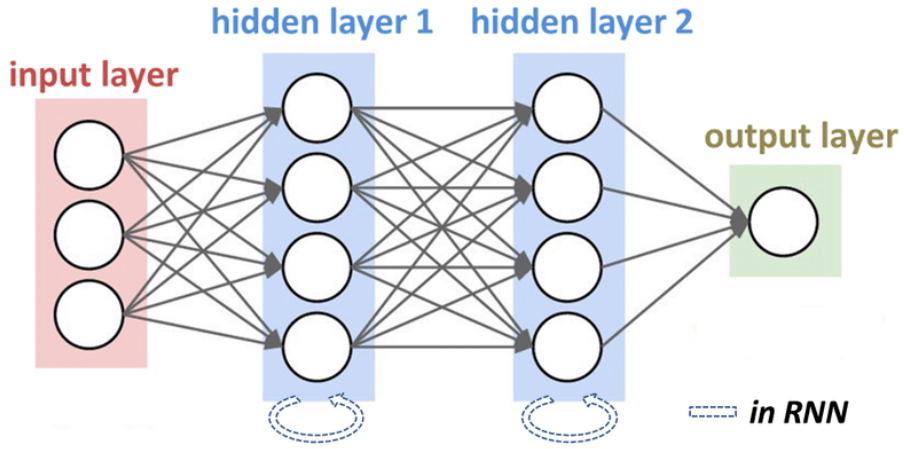
- **Recurrent Neural Networks**

Recurrent Neural Networks (RNNs) are a type of artificial neural network designed to work with sequential data and handle tasks where the order of information is crucial.

An RNN remembers every piece of information throughout time.

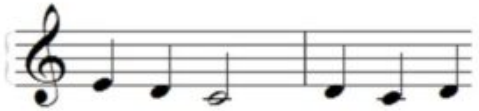
Apple's Siri and Google's voice search algorithm are exemplary applications of RNNs in machine learning.





Popular Architectures

- **Recurrent Neural Networks**

	x - input		y - output
Speech recognition		⟹	"Fuzzy Wuzzy was a bear. Fuzzy Wuzzy had no hair."
Music generation	∅	⟹	
Sentiment classification	"Decent effort. The plot could have been better."	⟹	
DNA sequence analysis	ACTGTACCCATGTGACTGCCC	⟹	ACTGTACCCATGTGACTGCCC
Machine translation	"El que no arriesga, no gana."	⟹	"If you don't take risks, you cannot win."
Video activity recognition		⟹	Running
Name entity recognition	"Ygritte says Jon Snow knows nothing."	⟹	"Ygritte says Jon Snow knows nothing."

Popular Architectures

- **Deep Belief Networks**

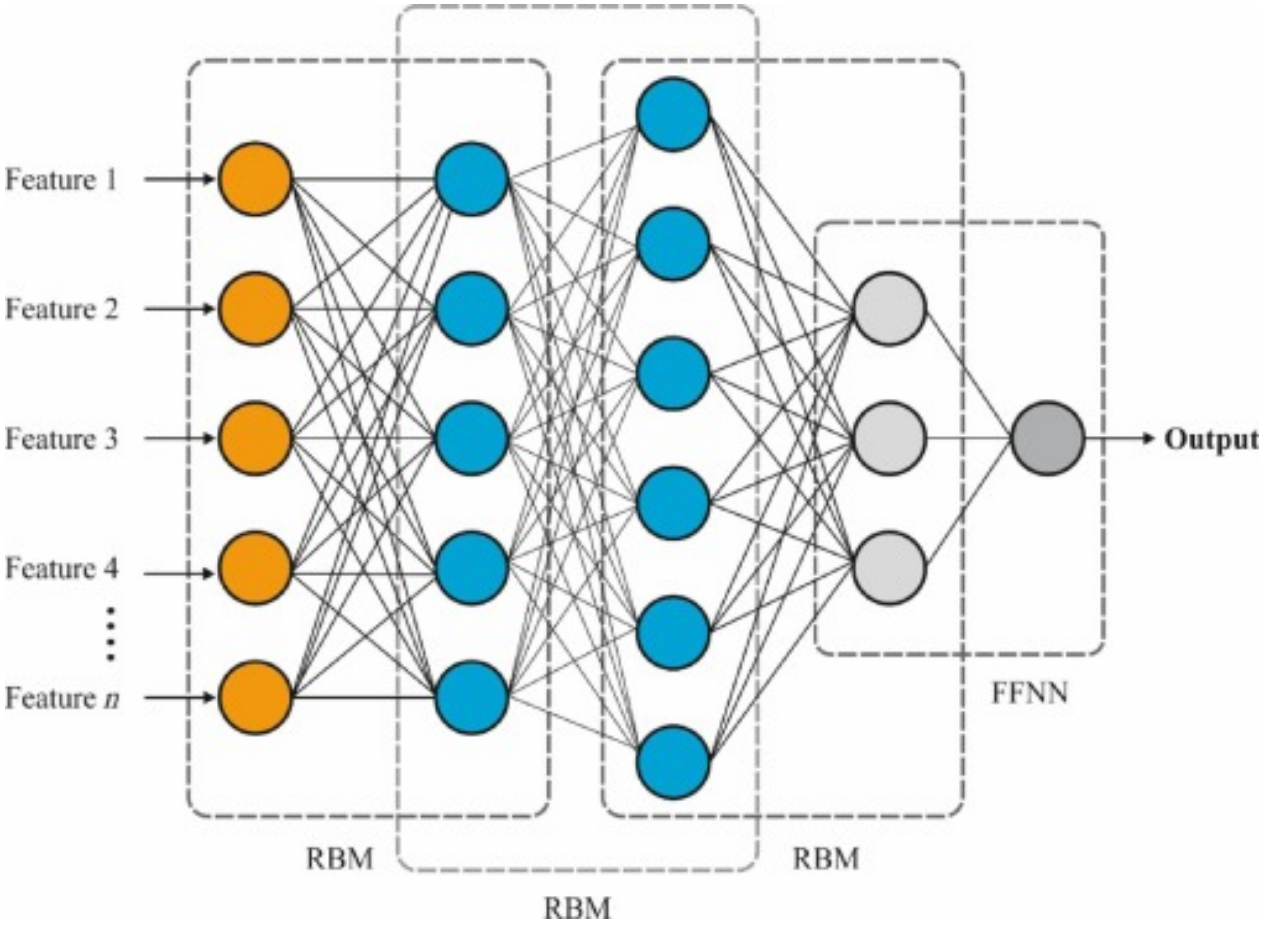
DBNs consist of multiple layers of stochastic(random probability), latent variables (often binary) and are structured as a stack of Restricted Boltzmann Machines (RBMs). Each layer captures progressively complex features of the input data.

DBNs are trained layer by layer in an unsupervised manner. The first layer is trained to capture basic patterns in the data, and subsequent layers are added and trained to capture higher-level representations.

The architecture of DBNs also makes them good at unsupervised learning, where the goal is to understand and label input data without explicit guidance. This characteristic is particularly useful in scenarios where labelled data is scarce or when the goal is to explore the structure of the data without any preconceived labels.

Popular Architectures

- **Deep Belief Networks**



Consider a DBN applied to image data.

The first layer might learn simple features like edges and corners, while deeper layers learn more complex features such as textures or object parts.

The final layer could represent high-level features like complete objects.

- **Graph Neural Networks**

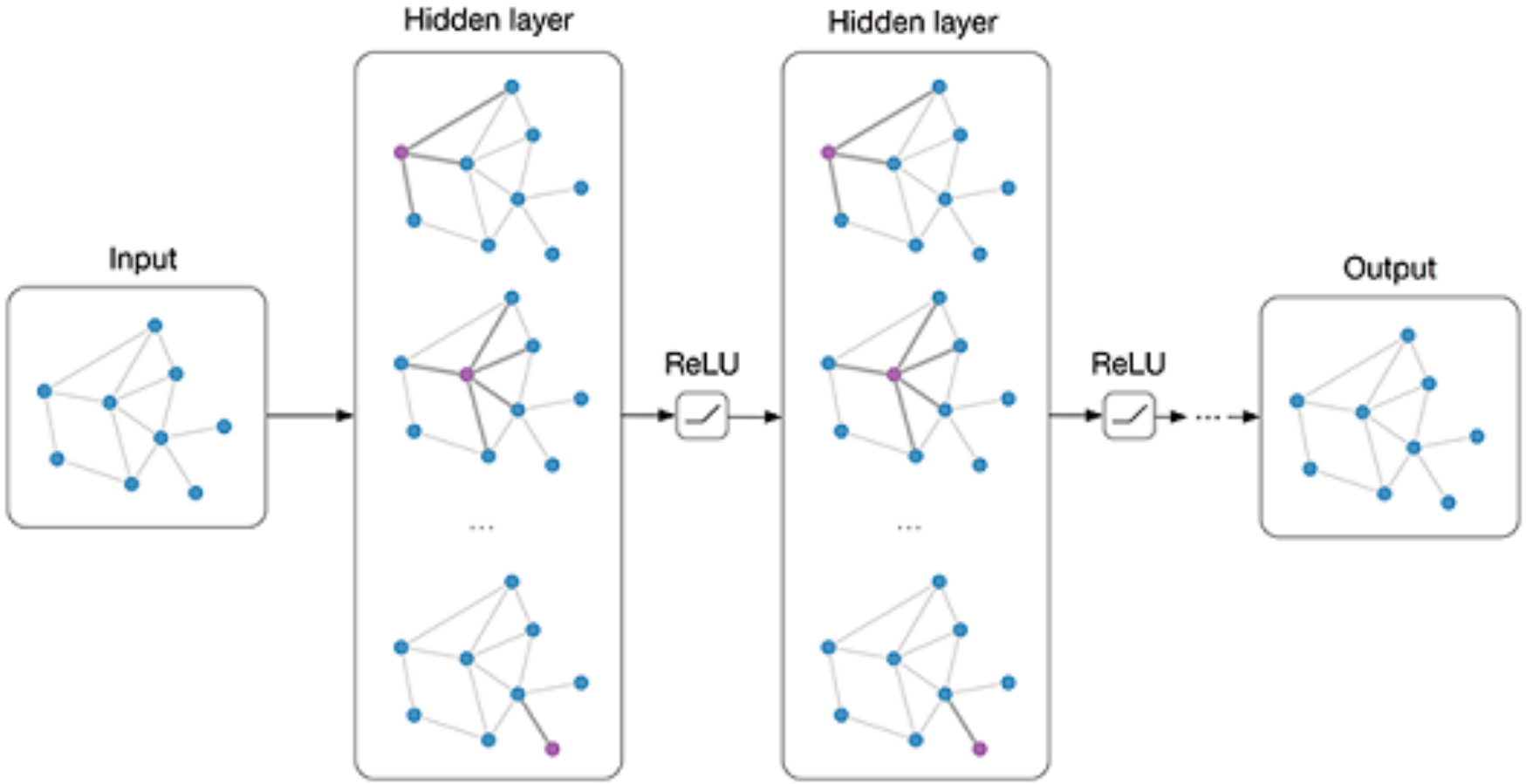
GNNs are a type of neural network architecture designed to process and analyze graph-structured data.

Unlike traditional neural networks that operate on grid-structured data like images or sequences, GNNs are tailored for data represented as graphs, which consist of nodes (vertices) and edges connecting these nodes.

Some examples of its usage include Recommender Systems, Social Network Analysis and more.

Popular Architectures

- **Graph Neural Networks**



Popular Architectures

- **Reinforcement Learning**

RL is a type of machine learning paradigm where an agent learns to make decisions by interacting with an environment.

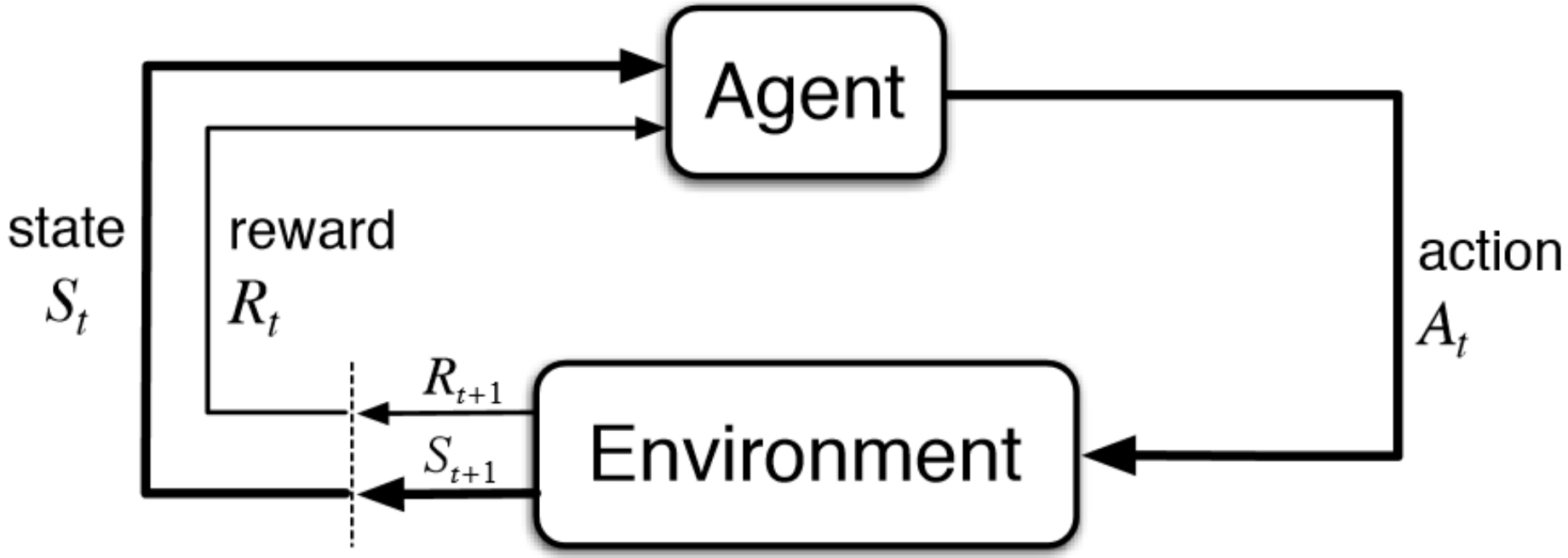
The goal of reinforcement learning is for the agent to learn a strategy that maximizes a cumulative reward over time.

The agent receives observations, takes actions, and receives rewards from the environment, aiming to learn a strategy that maximizes cumulative rewards over time.

RL is employed in scenarios where optimal decision-making is learned through trial and error in complex and dynamic environments.

Popular Architectures

- **Reinforcement Learning**



- **Generative Adversarial Networks**

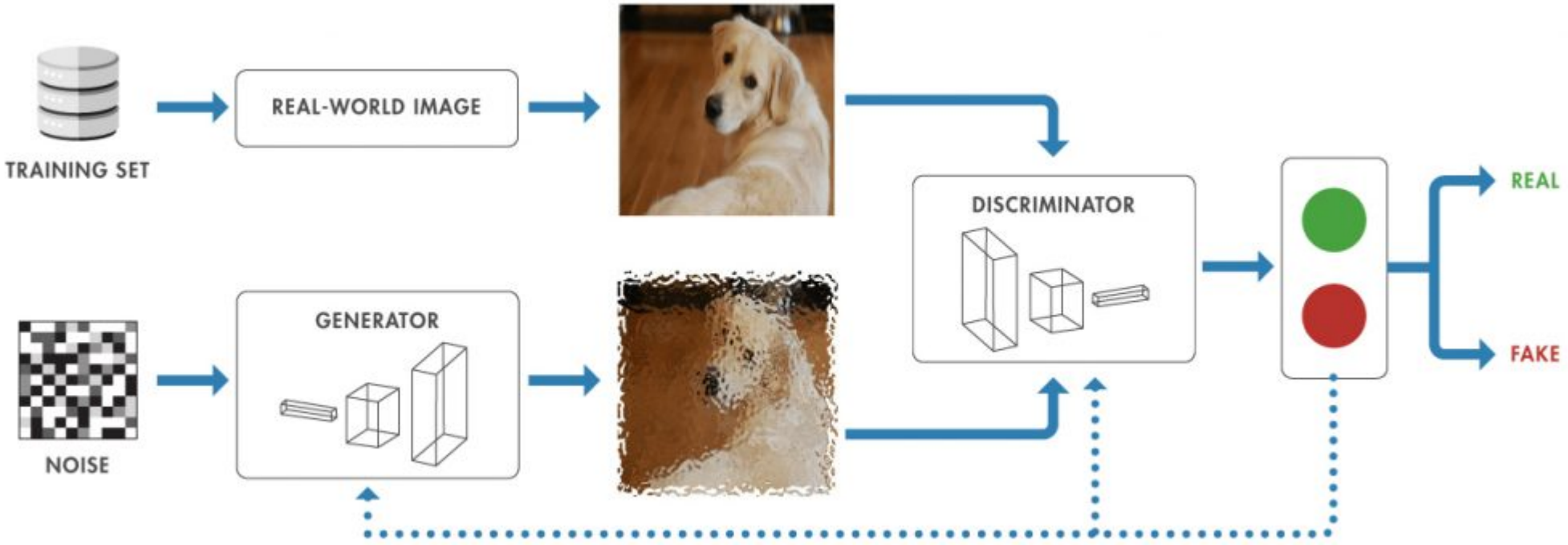
Generative Adversarial Networks (GANs) are a class of deep learning models composed of a generator and a discriminator trained in competition.

The generator creates synthetic data, while the discriminator evaluates its authenticity, leading to an adversarial learning process.

GANs have achieved remarkable success in generating realistic images, videos, and other data, with applications ranging from art creation to data augmentation.

Popular Architectures

- **Generative Adversarial Networks**



Popular Architectures

- **Transformers**

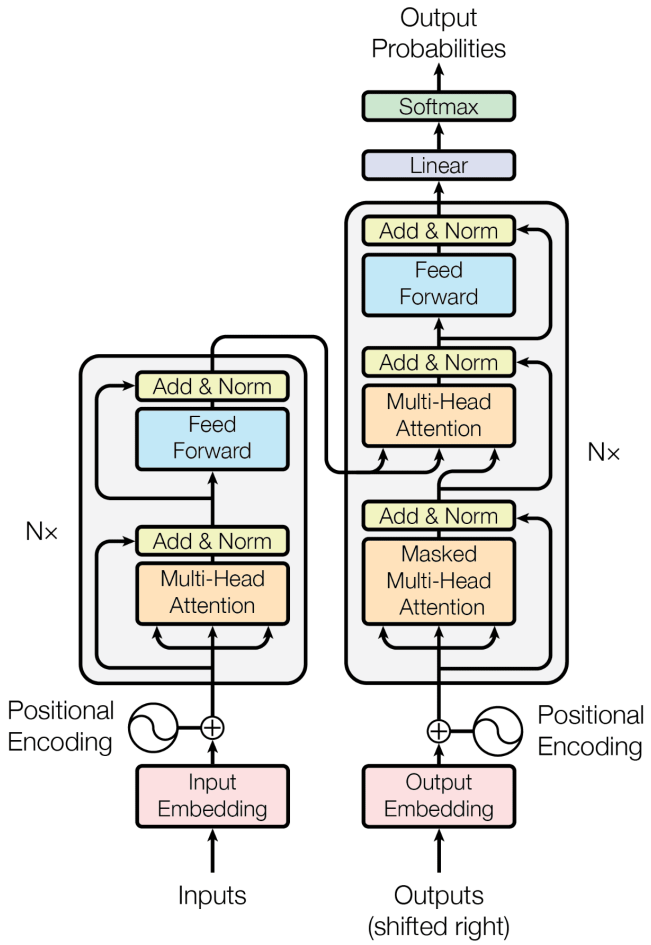
Transformers are a type of deep learning model architecture renowned for their self-attention mechanism, enabling the capture of intricate relationships within input sequences.

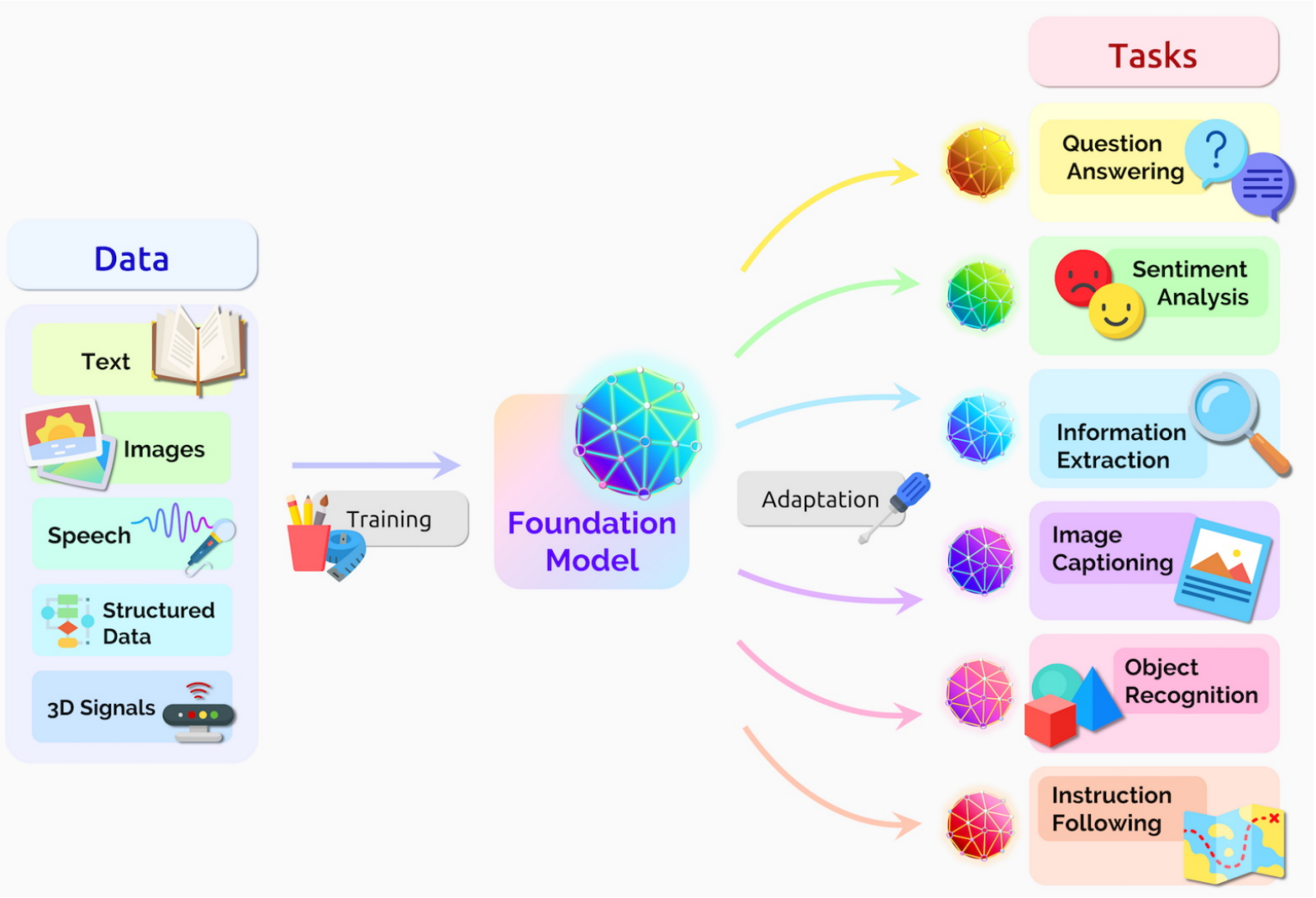
Their parallelization capabilities and encoder-decoder architecture make them efficient for various tasks, extending beyond natural language processing to computer vision.

Notably, models like BERT (Bidirectional Encoder Representations from Transformers), Google's neural network-based technique for natural language processing (NLP) pre-training have demonstrated exceptional performance.

Popular Architectures

- **Transformers**





Popular Architectures

- **GPT (Generative Pre-trained Transformer) Series:**

Models like ChatGPT, part of the GPT series, are transformer-based language models trained on diverse data, excelling in natural language generation and understanding.

- **Google Gemini**

Google Gemini is built with a decoder architecture featuring a 32k context length(words that the model can take into account when generating responses or predictions) and Multi Query Attention (MQA), Gemini is engineered for advanced contextual understanding, setting a new standard in AI architecture.

Popular Architectures

- **GPT4 vs Gemini:**

Capability	Benchmark Higher is better	Description	Gemini Ultra	GPT-4 API numbers calculated where reported numbers were missing
General	MMLU	Representation of questions in 57 subjects (incl. STEM, humanities, and others)	90.0% CoT@32*	86.4% 5-shot* (reported)
Reasoning	Big-Bench Hard	Diverse set of challenging tasks requiring multi-step reasoning	83.6% 3-shot	83.1% 3-shot (API)
	DROP	Reading comprehension (F1 Score)	82.4 Variable shots	80.9 3-shot (reported)
	HellaSwag	Commonsense reasoning for everyday tasks	87.8% 10-shot*	95.3% 10-shot* (reported)
Math	GSM8K	Basic arithmetic manipulations (incl. Grade School math problems)	94.4% maj1@32	92.0% 5-shot CoT (reported)
	MATH	Challenging math problems (incl. algebra, geometry, pre-calculus, and others)	53.2% 4-shot	52.9% 4-shot (API)
Code	HumanEval	Python code generation	74.4% 0-shot (IT)*	67.0% 0-shot* (reported)
	Natural2Code	Python code generation. New held out dataset HumanEval-like, not leaked on the web	74.9% 0-shot	73.9% 0-shot (API)

Acknowledgements & References

- https://www.geeksforgeeks.org/
- https://deeplearning.ai
- https://medium.com/@developer.yasir.pk/unveiling-googles-gemini-a-deep-dive-into-the-next-frontier-of-ai-ee41ffe90a9c

Thank You

Dr. Shylaja S S
Director of Cloud Computing & Big Data (CCBD), Centre for Data Sciences & Applied Machine Learning (CSSAML)
Department of Computer Science and Engineering
shylaja.sharath@pes.edu