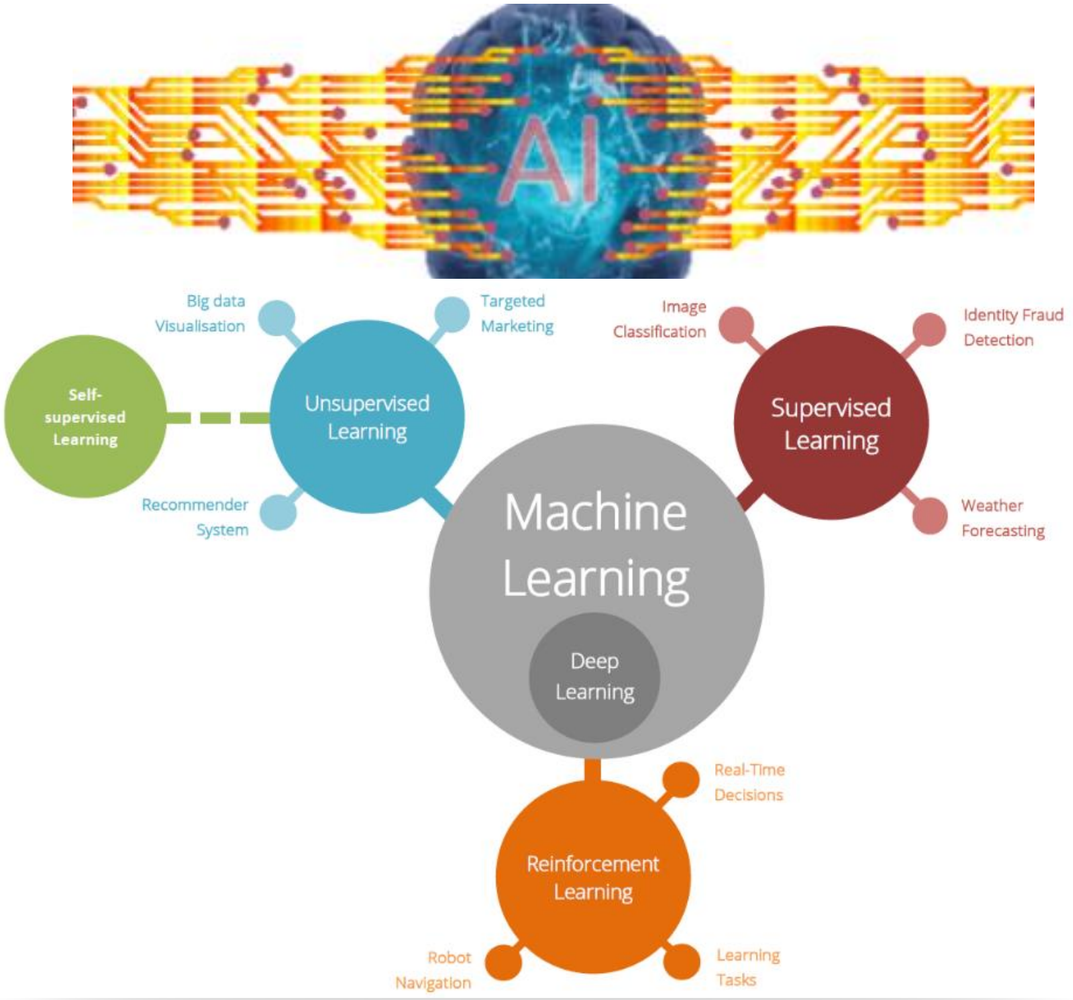
DroneX AI



Master AI Terminology

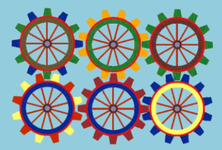
Terms of the Future!

Karn Singh

Key Terms

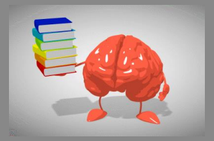
AI

AI is technology that enables computers and machines to simulate human learning, comprehension, problem solving, decision making, creativity and autonomy.



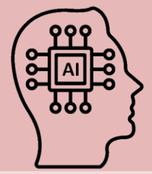
Machine Learning

Machine Learning is a subset of artificial intelligence that helps you build AI-driven applications.



Deep Learning

Deep Learning is a subset of machine learning that uses vast volumes of data and complex algorithms to train a model.



Generative AI

GAI refers to the use of AI to create new content, like text, images, music, audio, and videos. GAI is powered by foundation models (large AI models) that can multi-task and perform out-of-the-box tasks, including Q&A, summarization, classification and more.

LLM

Large language models (LLMs) are **AI systems capable of understanding and generating human language** by processing vast amounts of text data.

AGI

AGI refers to a type of AI that possesses the ability to understand, learn, and apply knowledge across a wide range of tasks, similar to human cognitive abilities.

ASI

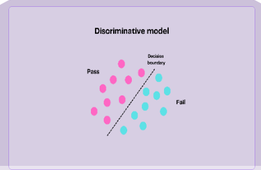
ASI is a hypothetical form of AI that surpasses human intelligence across all fields, including creativity, general wisdom, and social skills.

ANI

ANI also known as weak AI, is designed to perform specific tasks based on its programming and is not able to learn beyond its programmed capabilities.

Open AI

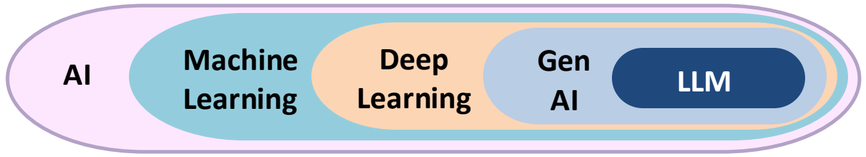
Open AI is an AI research company that aims to develop artificial general intelligence (AGI) in a way that benefits everyone.



Discriminative AI

Discriminative AI refers to a class of machine learning models that focus on classifying and predicting outcomes based on input data.

LLM Key Terms



Prompting

involves providing specific instructions or cues to guide the behavior of AI model. This can include asking the model to generate text, Q&A or perform tasks based on given input.

User Prompt

A user prompt is a specific input or request made by a user to an AI system. It serves as the starting point for the AI to generate a response or perform an action.

System Prompt

System prompting refers to the internal mechanisms or cues used by an AI system to elicit responses or actions. This can include pre-defined prompts that guide the AI's responses based on user interactions.

Meta Prompting

Meta Prompting is a Higher-level instructions that guide how prompts should be structured or interpreted.

Context Length

Maximum number of input words (tokens) a model can consider when generating an output

Tokenization

Process of breaking down text into smaller units (tokens), generally used as input for LLMs

Transformers

A popular LLM design known for its attention mechanism and parallel processing ability.

Attention

Popular mechanism in neural networks that allows focusing on specific parts of the input data.

Weight

Intrinsic parameters of a model that determine how it processes and generates text.

Embedding

A numerical format used to represent data features. Also called **vectors**.

LLM Learning Key Terms

In Context

It is a method where LLMs learn to perform tasks by being provided with examples within the prompt during inference time.

Zero-Shot Learning

It refers to the capability of an LLM to perform a task without having seen any specific examples of that task during training or prompting.

One-Shot Learning

One-shot learning algorithms can recognize and generalize patterns based on just one example of each class or category.

Few-Shot Learning

It is a Machine learning technique where a model learns to recognize patterns or make predictions from a small number of examples

Multi-Task Learning

Learning to perform multiple tasks & sharing knowledge between related tasks for better performance.

Meta Learning

Learning to learn by extracting general knowledge from diverse task and applying it to new ones.

Unsupervised Learning

Learning patterns and structures from data without specific guidance or labels

Reinforcement Learning

Learning through trial and error, with rewards or penalties based on generated outputs.

Adversarial Learning

Training against adversaries or competing models to improve robustness and performance.

Supervised Learning

Learning from labeled examples & associating inputs with correct outputs.

Federated Learning

Training across multiple decentralized devices without sharing raw data, preserving user privacy.

Active Learning

Interacting with human or the environment to select and label the most useful data for training.

LLM Fine-Tuning Key Terms

Fine Tuning

Fine-tuning is the process of further training a pre-trained machine learning model on a smaller, specialized dataset to adapt it for specific tasks or use cases.

PEFT

PEFT (Parameter-Efficient Fine-Tuning) is a technique for efficiently adapting large pre-trained models to downstream tasks by fine-tuning only a small subset of parameters

LoRA

LoRA is a Method to reduce the number of trainable parameters during fine-tuning by freezing all original model parameters and injecting a pair of rank decomposition matrices alongside the original weights.

RLHF

Reinforcement learning from human feedback (RLHF) is a technique that trains AI models by rewarding them based on human feedback to align the models with human preferences.

Reward Model

A reward model is an AI system that assigns numerical scores to outputs, representing human preferences.

Quantization

Quantization is the process of constraining an input from a continuous or otherwise large set of values, like real numbers, to a discrete set, such as integers.

GGUF

GGUF (GPT-Generated Unified Format) is a binary file format designed for efficiently storing and deploying large language models (LLMs), optimized for fast loading and saving of models while incorporating both model parameters and metadata.

Distillation

Distillation is a process in machine learning where a smaller, simpler model (student) is trained to replicate the behavior of a larger, more complex model (teacher) to improve efficiency while retaining performance.

PTQ

Post-Training Quantization (PTQ) is a technique that applies quantization to a pre-trained model after its training is complete, reducing the model's precision (e.g., from floating-point to integer) to decrease its memory

Pruning

Pruning is a technique in machine learning that involves removing parts of a model, such as nodes or branches in decision trees or neural networks, to reduce complexity and improve performance, particularly by preventing overfitting.

LLM and RAG Key Terms

RAG is an AI framework for improving the quality of LLM-generated responses by grounding the model on external sources of knowledge.

Knowledge Base

Collection of documents from which relevant information is retrieved in RAG

Chunking

The process of breaking the KB into smaller pieces for efficient storage and retrieval during RAG.

Indexing

Organizing and storing KB chunks in a structured manner for efficient retrieval.

Vector Database

A database optimized for storing and retrieving vector representations generated from the KB.

Vector Search

The process of finding the most relevant KB chunks based on vector similarity scores for a given input query.

Retrieval

The approach used to rank and fetch KB chunks from the vector search. These chunks serve as additional context for the LLM.

LLM Agent

LLM applications that can execute complex tasks by combining LLMs with modules like planning and memory.

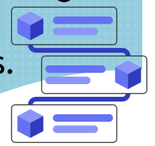
Agent Memory & Planning

Agent Memory a module that stores the agent's past experiences and interactions with the user and environment.

Agent Planning a module that divides the agent's tasks into smaller steps to address the user's request efficiently.

CoT

CoT (Chain of Thought) is a prompting technique that encourages language models to articulate their reasoning process step-by-step when solving problems, leading to more accurate and interpretable results.



ReAct

ReAct is a framework that enables AI agents to effectively reason about tasks and take actions in interactive environments by combining decision-making processes with real-time feedback loops.

Function Calling : The ability of LLM agents to request information from external tools and APIs to execute a task.

LLM Vulnerabilities and Attack Key Terms

Alignment

Ensuring that the behavior of an LLM is consistent with human values.

Watermarking

Embedding hidden markers into LLM-generated content to track its origin or authenticity.

Vulnerability

Weaknesses or flaws in LLMs that can be exploited for malicious purposes.

Robustness

The ability of an LLM to perform accurately despite encountering adversarial inputs.

Adversarial Attacks

Deliberate attempts to trick LLMs with carefully crafted inputs, causing them to make mistakes.

Black-Box Attacks

Trying to attack an LLM without knowing its internal workings or parameters.

White-Box Attack

Attacking an LLM with full knowledge of its internal architecture and parameters.

Deep-Fakes

Synthetic media generated by LLMs, often used to create realistic but fake images or videos.

Jailbreaking

Attempting to bypass security measures around an LLM to make it produce unsafe outputs.

Prompt Injection

Hijacking the LLM's original prompts to make it perform unintended tasks.

Prompt Leaking

Tricking an LLM to reveal information from its training or inner workings.

Red-Teaming

Assessing the security and robustness of LLMs through simulated adversarial attacks.

LLM Terms You Should Know!

Hallucination

Tendency of LLMS to sometimes generate incorrect or non factual information

Self Supervision

Training method that involves the model learning from data without clear labels often by predicting hidden parts of the input.

Responsible AI

Ensuring that AI systems are developed and used ethically, fairly, and transparently.

MoE

Mixture of Experts A machine learning technique where multiple expert networks are used to divide a problem space into homogeneous regions.

MM LLM

Multimodel LLM is capable of understanding & generating content across multiple modalities images and audio along with text.

Domain Adaptation

Adapting a model trained on one domain to perform well in a different but related domain.

Foundation LLM

LLM design to generated and understand human like text across a wide range of use cases.

GDPR

General Data Protection Regulation A regulation in the EU that protects individuals' privacy rights and governs data handling.

XAI

Explainable AI The concept of making the AI model's outputs understandable and transparent to humans.

LLM Ops

Managing and optimizing operations for LLM deployment and maintenance.

PII

Personally Identifiable Information Info that can identify an individual. It should not be stored or used without proper processes and user consent.

LLM Bias

Refers to systematic prejudices in the LLM's predictions, often stemming from biased training data.

DroneX AI

REPOST

WAS THIS POST USEFUL?

FOLLOW FOR MORE!