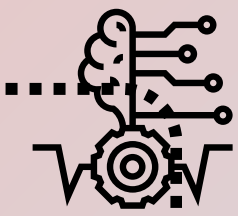
LLM Lingo: Must-Know Terms

Around 70 of the most frequently used LLM terms, along with simple explanations, sorted by topics.

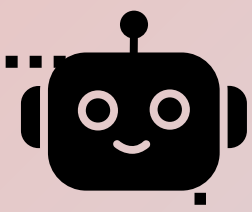
Part 1
Basic LLM Related Terms

Basic terms that are in the LLM space like **foundation model, prompting, context-length** etc.



Part 2
Fine-tuning

Fine-tuning related terms like **PEFT, quantization, LoRA** etc.



Part 3
RAG and Agents

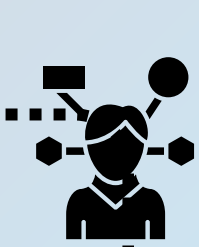
Terms used in RAG like **vector databases, chunking** etc. and some basic agent terms like **function calling**

Part 4
Enterprise Ready LLMs

Terms that are commonly used while deploying LLMs like **compliance, GDPR, PII, ethics** etc.

Part 5
LLM Vulnerabilities & Attacks

Terms related to LLM attacks like **adversarial attacks, prompt injection, prompt leaking** etc.



Part 6
LLM Learning Paradigms

Different learning paradigms that LLM use during training like **unsupervised learning, reinforcement learning** etc

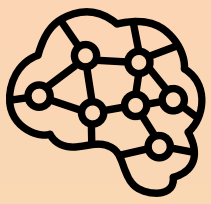
Created By: Aishwarya Naresh Reganti

LLM Lingo: Must-Know Terms

Part 1

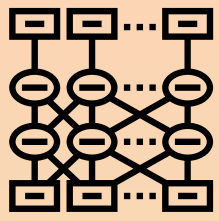
Created By: Aishwarya Naresh Reganti

Foundation Model



LLM designed to generate and understand human-like text across a wide range of use-cases

Transformer



A popular LLM design known for its attention mechanism and parparallel processing abilities

Prompting

Providing carefully crafted inputs to an LLM to generate desired outputs

Context-Length

Maximum number of input words/tokens an LLM can consider when generating an output.

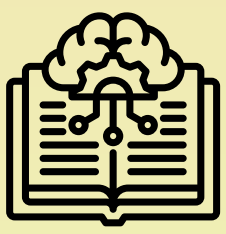
Few-Shot Learning

Providing very few examples to an LLM to assist it in performing a specific task.

Zero-Shot Learning

Providing only task instructions to the LLM relying solely on its pre-existing knowledge

RAG



Retrieval-Augmented Generation. Appending retrieved information to improve LLM response

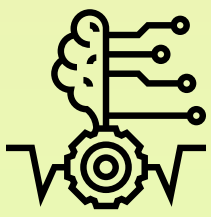
Knowledge Base(KB)

Collection of documents from which relevant information is retrieved in RAG

Vector Database

Stores vector representations of the KB, aiding the retrieval of relevant information in RAG.

Fine-Tuning



Adapting an LLM to a specific task or domain by further training it on task-specific data.

Instruction Tuning

Adjusting an LLM's behavior during fine-tuning by providing specific guidelines/directives

Hallucination

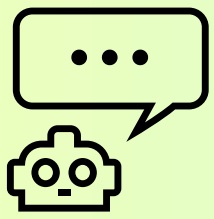
Tendency of LLMs to sometimes generate incorrect or non-factual information.

LLM Lingo: Must-Know Terms

Part 2: Fine-Tuning Edition

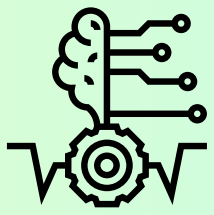
Created By: Aishwarya Naresh Reganti

In-Context Learning



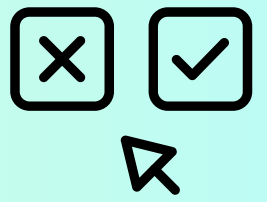
Integrating task examples into prompts, enabling LLMs to handle new tasks without fine-tuning.

SFT



Supervised Fine-Tuning. Updating a pre-trained LLM with labeled data to perform a specific task.

Contrastive Learning

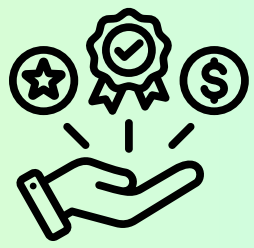


Fine-tuning method that improves LLM by teaching it to discern data similarity and differences.

Transfer Learning

Applying pre-trained knowledge from large datasets to improve LLM performance on smaller, task specific data.

Reward Modeling

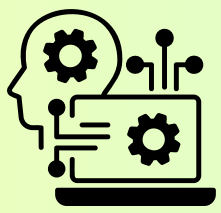


Designing objectives to reward LLM outputs during the reinforcement learning process.

Reinforcement Learning

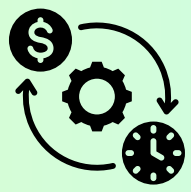
Training LLMs through trial and error, with rewards/penalties based on its generated outputs

RLHF



Reinforcement Learning from Human Feedback. Human feedback is used as reward/penalty for LLM

PEFT

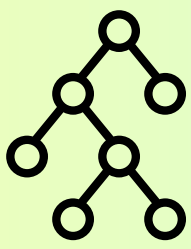


Parameter-Efficient Fine-Tuning updates only few parameters of LLMs and is hence both compute and cost efficient.

Quantization

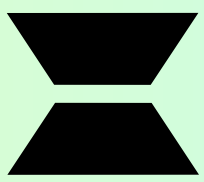
Reducing the precision of LLM parameters to save computational resources without sacrificing performance.

Pruning



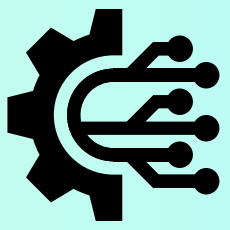
Trimming surplus connections or parameters to make LLMs smaller and faster yet performant

LoRA



Low-Rank Adaption is a PEFT method that inserts a smaller set of new weights to the LLM & trains only those.

Freeze Tuning



Fine-tune with most of the LLM's weights frozen, except for some layers, generally, the task specific layers

LLM Lingo: Must-Know Terms

Part 3: RAG + LLM Agents

Created By: Aishwarya Naresh Reganti

RAG or Retrieval-Augmented Generation in LLMs, combines regular language generation with information retrieval, enabling models to access external knowledge sources and improve the quality and relevance of their generated outputs. Below are some key terms:

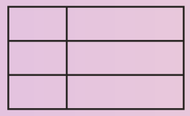
Knowledge Base(KB)

Collection of documents from which relevant information is retrieved in RAG

Chunking

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

Embedding Model

An LLM that converts KB text chunks into numerical format called vectors/embeddings

Vector Database

Database optimized for storing and retrieving vector representations generated from the KB

Vector Search

Finding the most relevant KB chunks based on vector similarity scores for a given input query.

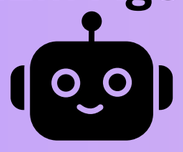
Retrieval

Approach used to rank and fetch KB chunks from the vector search. This will serve as additional context for the LLM.

AGI

Artificial General Intelligence aims to create machines that can learn and reason like humans across various tasks.

LLM Agent



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

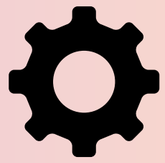
Agent Memory

A module that stores the agent's past experiences and interactions with the user and environment.

Agent Planning

Module that divides the agent's tasks into smaller steps to address the user's request efficiently.

Function Calling



Ability of LLM agents to request information from external tools and APIs in order to execute a task

LLM Lingo: Must-Know Terms

Part 4: Enterprise Ready LLMs

Created By: Aishwarya Naresh Reganti

LLM Bias

Systematic prejudices in the LLM's predictions, often stemming from training data

XAI

Explainable AI. Making the model’s outputs understandable and transparent to humans.

Responsible AI

Ensuring ethical, fair, and transparent development and use of AI systems.

AI Governance

Legal policies & and frameworks that regulate the development & deployment of AI systems

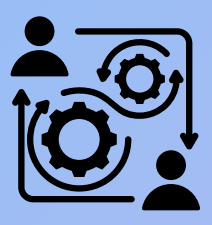
Compliance

Ensuring adherence to legal requirements in the development & deployment of AI systems.

GDPR

General Data Protection Regulation protecting individuals' privacy rights and governing data handling in the EU

Alignment



Ensuring that the outputs of LLMs are consistent with human values and intentions.

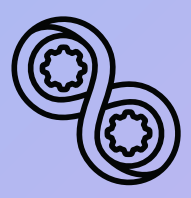
Model Ethics

Ensuring ethical behavior (transparency, fairness, accountability etc.) when deploying public facing AI

PII

Personally Identifiable Information. Should not be stored or used without proper processes and user consent.

LLMOps



Managing and optimizing operations for LLM deployment and maintenance.

Privacy-preserving AI

Methods to train and use LLMs while safeguarding sensitive data privacy.

Adversarial Defense

Methods to prevent malicious attempts to manipulate LLMs, ensuring their security.

LLM Lingo: Must-Know Terms

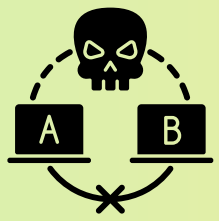
Part 5: LLM Vulnerabilities and Attacks

Created By: Aishwarya Naresh Reganti

Adversarial Attacks

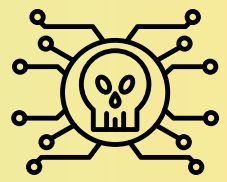
Deliberate attempts to trick LLMs with carefully crafted inputs, causing them to make mistakes.

Black-Box Attacks



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

White-Box Attacks



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

Vulnerability

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

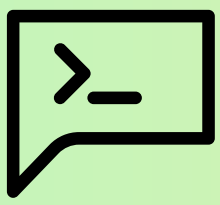
Deep-fakes

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

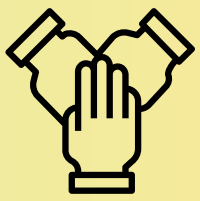


Hijacking the LLM's original prompts to make it perform unintended tasks

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.

Robustness

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

Alignment

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

Watermarking

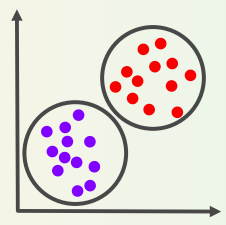
Embedding hidden markers into LLM-generated content to track its origin or authenticity.

LLM Lingo: Must-Know Terms

Part 6: LLM Learning Paradigms

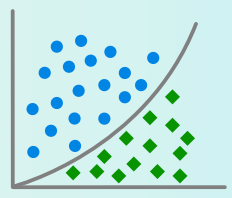
Created By: Aishwarya Naresh Reganti

Unsupervised Learning



Learning patterns and structures from data without specific guidance or labels.

Supervised Learning



Learning from labeled examples & associating inputs with correct outputs.

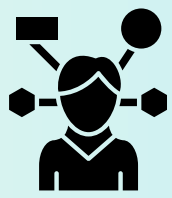
Reinforcement Learning

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

Meta-Learning

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

Multi-task Learning



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

Zero-Shot Learning

Providing only task instructions to the LLM relying solely on its pre-existing knowledge

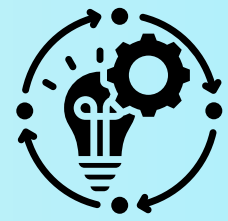
Few-Shot Learning

Learning from a small number of examples for new tasks and adapting quickly with minimal data.

Online Learning

Continuously learning from incoming data streams and updating knowledge in real-time.

Continual Learning



Learning sequentially from a stream of tasks or data without forgetting previously learned knowledge.

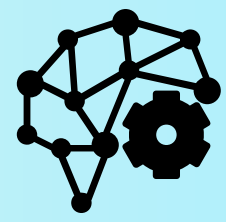
Federated Learning

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

Adversarial Learning

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

Active Learning



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