

Large Language Models (LLMs)

BY: Sarwat Majeed

(Student Leader)

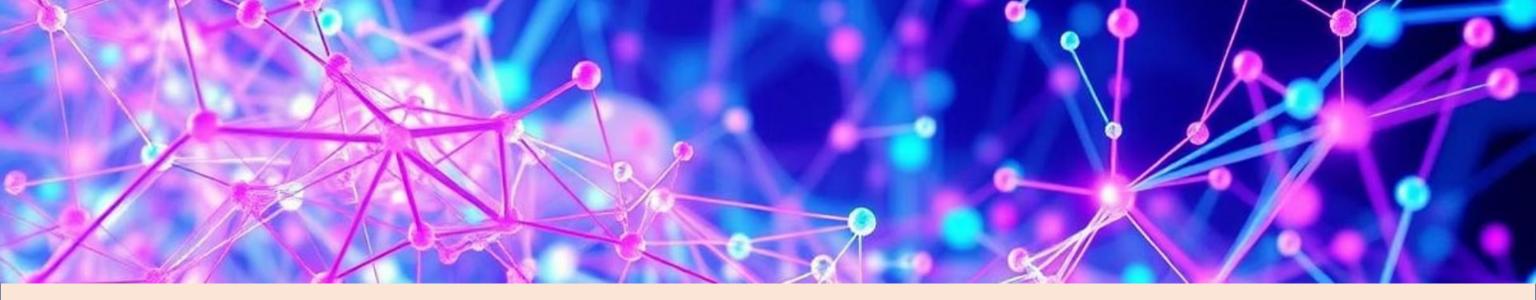
Roll #: 00331141

Slot: Wednesday 7 to 10



Understanding Large Language Models (LLMs)

Large Language Models (LLMs) generate human-like text by training on vast datasets. Examples include GPT-4, Bard, and LLaMA. They are revolutionizing industries with their versatile capabilities.



What Are Large Language Models?

Definition

Deep learning models with billions of parameters enable language understanding.

Training Data

Trained on internet text, books, code, and diverse data sources.

Scale

Examples: GPT-3 (175B), PaLM (540B), GPT-4 (over 1 trillion).

Foundation Models

Adaptable for many different language tasks across industries.



How LLMs Work: Training Phase

Self-Supervised Learning

The model predicts the next word in sequences without explicit labels.

Loss Function

Measures prediction errors to guide model improvement.

Backpropagation

Adjusts billions of parameters to minimize prediction errors.

Compute Power

4

Needs thousands of GPUs or TPUs over weeks or months of training.

How LLMs Work: Inference Phase

Prompt Engineering

Creating precise input queries to guide model responses effectively.

Tokenization

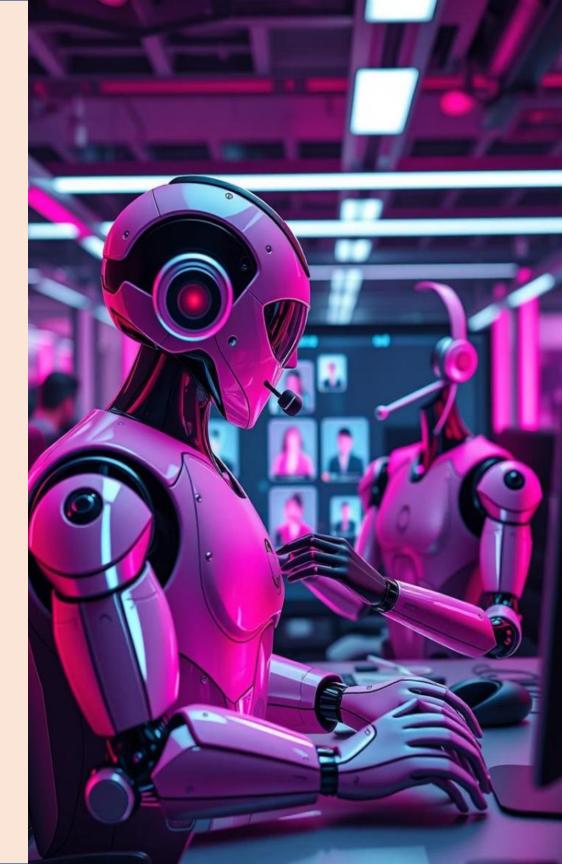
Breaking text into smaller pieces processed by the model.

Attention Mechanism

Focuses on relevant input parts for context-aware generation.

Decoding

Generates text based on learned probabilities and context clues.





Use Case: Content Generation

Marketing Content

LLMs create blogs, ads, and social media updates fast.

Jasper.ai

Generates marketing copy, valued at 1.5B.

Copy.ai

Automates writing with 13.9M funding.

Efficiency

Produces ten times more content than humans.

Use Case: Customer Service

Automation

LLMs power chatbots and virtual assistants for fast support.

Intercom

Uses AI to automate client interaction, valued at 1.6B.

Ada Support

AI-driven service automation with 130M funding.

Cost Reduction

Can cut support costs by up to 70%.



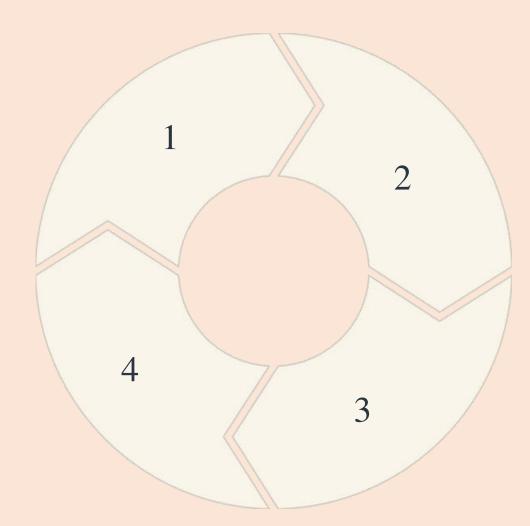
Use Case: Code Generation

Writing Assistance

LLMs help draft and autocomplete code efficiently.

Time Savings

Reduces coding time by up to 40%.



GitHub Copilot

AI pair programmer powered by OpenAI Codex.

Tabnine

Popular AI code completion tool with 1 million users.

Conclusion: The Future of LLMs

Advancements Models will grow larger, faster, and more efficient. Integration Embedded in diverse industries and daily apps. Ethics 3 Focus on responsible AI development and fairness. Impact Changing how humans interact with technology forever.