



Large Language Models

[Source: IE](#)

Why in News?

In the era of advanced [artificial intelligence \(AI\)](#), the emergence of [Large Language Models \(LLMs\)](#) has revolutionized the way computers interact with humans and process language. From enhancing virtual conversations to powering creative tasks, LLMs have paved the way for a new frontier in the realm of AI technology.

What are Large Language Models (LLMs)?

- **Definition:**
 - LLMs are large general-purpose language models capable of solving common language problems such as text classification, question answering, and text generation.
 - These models are **trained on massive datasets to understand patterns**, structures, and relationships within human language.
- **Types of Large Language Models (LLMs)**
 - **Based on Architecture:**
 - **Autoregressive Models:** Predict the **next word in a sequence** based on previous words. **Example: GPT-3.**
 - **Transformer-based Models:** Utilise a specific **artificial neural network architecture** for language processing. Examples: **LaMDA, Gemini (formerly Bard).**
 - **Encoder-decoder Models:** Encode input text into a representation and then decode it into another language or format.
 - **Based on Training Data:**
 - **Pretrained and Fine-tuned Models:** Adapt to specific tasks through fine-tuning on particular datasets.
 - **Multilingual Models:** Capable of understanding and generating text in multiple languages.
 - **Domain-specific Models:** Trained on data related to specific domains like legal, finance, or healthcare.
 - **Based on Size and Availability:**
 - **Size:** Large models require more computational resources but offer better performance.
 - **Availability:** Open-source models are freely available, while closed-source models are proprietary.
 - **Examples of open-source LLMs:** LLaMA2, BLOOM, Google BERT, Falcon 180B, OPT-175 B.
 - **Examples of closed-source LLMs:** [GPT 3.5 by OpenAI](#), [Gemini by Google](#).
- **Operational Mechanisms of LLMs:**
 - At their core, LLMs utilize [deep learning techniques](#), to **predict the probability of words or sequences** given preceding text.
 - LLMs analyze patterns and relationships in data to predict the next word or sequence based on input prompts, akin to **how humans comprehend language**.
 - LLMs typically rely on [transformer models](#), such as the [Generative Pre-trained](#)

[Transformer \(GPT\)](#), with attention mechanisms for contextual understanding.

▪ **Applications of LLMs:**

- LLMs generate **human-like content**, from stories to songs, and act as virtual assistants, excelling in sentiment analysis, translation, and text summarization, crucial for marketing strategies.

▪ **Advantages of LLMs:**

- LLMs can adapt to various tasks and domains, leveraging their extensive training data to generalise patterns.
- They can perform well even with limited domain-specific data, thanks to their ability to learn from general language training datasets.
- As more data and parameters are added, LLMs **continuously enhance their performance**, making them valuable assets in evolving AI landscapes.

What are Large Action Models (LAMs)?

- LAMs are **specialized AI models** built to perform specific **tasks or sequences of actions**, often beyond just understanding and generating text.
 - **LAMs** can understand human intention and predict actions. LAMs are designed to help with repetitive tasks.
- They are designed to execute **actions based on inputs, which may include text, images, or other forms of data.**
- LAMs can be used in various applications such as **virtual assistants, robotic systems, automated customer service**, and more.
 - **Example of LAM:** [Rabbit r1](#).
- These models are trained on **datasets that include both linguistic information and action-oriented data** to learn how to perform tasks based on given contexts.

UPSC Civil Services Examination, Previous Year Questions (PYQs)

Prelims

Q. With the present state of development, Artificial Intelligence can effectively do which of the following? (2020)

1. Bring down electricity consumption in industrial units
2. Create meaningful short stories and songs
3. Disease diagnosis
4. Text-to-Speech Conversion
5. Wireless transmission of electrical energy

Select the correct answer using the code given below:

- (a) 1, 2, 3 and 5 only
- (b) 1, 3 and 4 only
- (c) 2, 4 and 5 only
- (d) 1, 2, 3, 4 and 5

Ans: (b)

Q 2. "The emergence of the Fourth Industrial Revolution (Digital Revolution) has initiated e-Governance as an integral part of government". Discuss. (2020)

