Large Language Model

Why in news?

The ability of Generative AI models to “converse” with humans and predict the next word or sentence is due to Large Language Model, or LLM.

What is Large Language Model?

- They are large general-purpose language models that can be pre-trained and then fine-tuned for specific purposes.
- These models are trained to *solve common language problems* such as text classification, question answering, text generation across industries, document summarisation, etc.
- **Deep learning techniques-** It involves the training of artificial neural networks, which are mathematical models which are believed to be inspired by the structure and functions of the human brain.
- The neural network learns to predict the probability of a word or sequence of words given the previous words in a sentence.
- **Next word prediction engines**- Once trained, an LLM can predict the most likely next word or sequence of words based on inputs also known as prompts.
- **Features**- LLMs has primary features like large, parameter and general purpose
- **Large**- This means
    - The enormous size of training data
    - **High parameter count-** It encode the knowledge and memories the model acquires during training.
- **General purpose**- It means the model is sufficient to solve general problems that are based on the commonality of human language regardless of specific tasks, and resource restrictions.
- **Parameter**- It play a crucial role in determining the model’s performance, it defines the model’s ability to solve specific task

Types of LLMs
• **Autoregressive model-** GPT 3 is an example of this model as they predict the next word in a sequence based on previous words. • **Transformer model-** LaMDA or Gemini are based on this as they use specific type of neural network architecture for language processing. • **Encoder-decoder model**- It encode input text into a representation and then decode it into another language or format.

What are the application of LLMs?

- **Natural Language Understanding (NLU)-** LLMs power advanced chat bots capable

of engaging in natural conversations which can be ued They can be used to create intelligent virtual assistants for tasks like scheduling, reminders, and information retrieval.

- **Content generation-** Creating human-like text for various purposes, including content creation, creative writing, and storytelling.
- **Language translation-** They can aid in translating text between different languages with improved accuracy and fluency.
- **Text summarization-** Generate concise summaries of longer texts or articles.
- **Sentiment analysis**- It helps in analysing and understanding sentiments expressed in social media posts, reviews, and comments.
- **Legal document analysis-** It can be applied in the legal domain for document analysis, contract review, and legal research, improving efficiency in legal processes.
- **Education-** It can be integrated into educational tools for tasks such as automated grading, generating educational content, and providing language learning support.
- **Code Generation**: LLMs have been utilized for generating code snippets and assisting developers in writing software by understanding and interpreting programming languages.
- **Question answering systems**- LLMs can be trained to provide accurate and contextually relevant answers to user queries, making them valuable for information retrieval applications.
- **Accessibility features-** It enhance accessibility by providing text-to-speech or speech-to-text capabilities, helping individuals with visual or hearing impairments.
- **Medical information**- It can assist in extracting relevant information from medical literature, aiding researchers and healthcare professionals in staying updated with the latest advancements.
- **Conversational AI**- It can serve as the backbone for chatbots and virtual assistants, providing natural and contextually aware interactions in customer support, information retrieval, and other conversational interfaces.

Advantages	Disadvantages
Zero shot learning- LLMs can generalize to tasks they were not explicitly trained for, showcasing adaptability to new applications	**High cost**- Setting up the computing power for large models requires significant investment.
Efficient data handling- It can process vast amounts of data, making them suitable for tasks like language translation and document summarization.	**Data availability**- Obtaining a large, high-quality text corpus can be challenging
Fine tuning- LLMs can be fine-tuned on specific datasets or domains, enabling continuous learning and adaptation.	**Bias**-Many large data sets used for training LLMs contain biases and prejudices leading to biased or discriminatory content.
Smooth training- LLMs streamline training by leveraging unlabelled data, and accelerates the process which saves time and resource	**Time consuming**- It takes months of training and human fine-tuning are necessary for optimal performance.
Automation- They can automate language-related tasks, freeing human resources for more strategic aspects of projects.	**Environmental impact**- Training LLMs contributes to carbon emissions.

Performance- It provide fast responses, improving overall business efficiency and productivity, their high-performance capabilities enhance language-related tasks and content delivery	**Hallucination-** LLMs may generate incorrect content without relying on learned data, this may lead to lack of accuracy or validity.

Reference

- Indian Express- LLMs are backbone of AI chatbots

