A Large Language Model (LLM) is a type of artificial intelligence model designed to understand and generate human-like text. These models are trained on vast amounts of text data—ranging from books and articles to websites—to learn patterns in language, grammar, facts, and reasoning. The term "large" refers to the size of the model in terms of parameters (the components that learn from data) and the volume of training data used. Popular LLMs include OpenAI's GPT (Generative Pre-trained Transformer) series, Google's Gemini, Meta's LLaMA, and Anthropic's Claude.

LLMs are based on the transformer architecture, which enables them to handle long-range dependencies in text and perform complex tasks such as translation, summarization, question answering, and even creative writing. After pretraining, these models can be fine-tuned for specific domains or applications. The key innovation of LLMs is their ability to generalize across a wide range of tasks without needing task-specific programming—this is what makes them so powerful and widely applicable.

However, LLMs also come with challenges. They can produce incorrect or misleading information (known as "hallucinations"), are data-hungry and computationally expensive to train, and raise ethical concerns around bias, misinformation, and privacy. Despite these limitations, LLMs have become central to modern AI, powering applications in customer support, education, coding, research, and more. As the technology continues to evolve, so do discussions around how to use it responsibly and safely.