GEN AI

Input from User → Model → gives prediction as output

← ML Model

Language as input → Model → predicts next best word

← General Language Model

Sequence of words as input → LLM → predicts sequence of words

Require large scale computation

Models are very complex

ex: Cost around ₹20 crores INR

Large no. of parameters are used (in billions etc)

↑ Whole internet of data has been used (websites, dictionaries, news, blogs, social media, opensource datasets ...)

trained on enormous amount of data

ex: GPT released in 2018 (110 million parameters)
GPT 3 2021 (175 billion)

Some well known LLMs

- Open AI
- Gemini
- Grok
- DeepSeek
- Qwen (by Alibaba)
- ...

How to use LLMs

- General User → LLM interface provided by Company (Open AI - chat GPT)
- Using an API - Developer