

**An LLM stands for Large Language Model.** It's a type of artificial intelligence model designed to understand and generate human-like text. Here's a clear breakdown:

---

## 1. Definition

- **LLM = Large Language Model**
  - A model trained on **massive amounts of text data** to predict and generate text.
  - Examples: **GPT-4, LLaMA 3, Claude, Bard**, etc.
- 

## 2. Key Capabilities

1. **Text generation** – Write essays, stories, or code.
2. **Conversation** – Chat and answer questions.
3. **Summarization** – Condense long documents into short summaries.
4. **Translation** – Translate between languages.
5. **Reasoning** – Solve problems or answer questions logically

**LLaMA 3** is a **large language model (LLM)** developed by **Meta (formerly Facebook)**. It's the third generation of the LLaMA series, designed for research and production applications. Here's a clear breakdown:

---

## Key Features of LLaMA 3

1. **Purpose**
  - Designed to generate human-like text.
  - Can answer questions, summarize text, write code, translate languages, and more.

## How to use it on your local windows pc

Step 1:-

Download Ollama and install it form

<https://ollama.com/>

step 2:- open command prompt and run :-

**ollama pull llama3**

```
C:\Users\User>ollama pull llama3
pulling manifest
pulling 6a0746a1ec1a: 100%
pulling 4fa551d4f938: 100%
pulling 8ab4849b038c: 100%
pulling 577073ffcc6c: 100%
pulling 3f8eb4da87fa: 100%
verifying sha256 digest
writing manifest
success
```

And then run it :-

```
C:\Users\User>ollama run llama3
>>> what is captial of india
The capital of India is New Delhi.
>>> Send a message (/? for help)
```

## Install required packages

```
pip install gradio langchain langchain-ollama
```

**simplechat.py :-**

```
from langchain_ollama.llms import OllamaLLM

# Initialize
llm = OllamaLLM(model="llama3", temperature=0.7)

print("Chat with LLaMA 3 (type 'exit' to quit)\n")
while True:
    user_input = input("You: ")
    if user_input.lower() in ["exit", "quit"]:
        break

    response = llm.invoke(user_input)
```

```
print("LLaMA 3:", response)
```

and chat-gradio.py file code:-

```
import gradio as gr
from langchain_ollama.llms import OllamaLLM

# Initialize LLaMA 3
llm = OllamaLLM(model="llama3", temperature=0.7)

# Function to handle chat
def chat_with_llama(user_input):
    response = llm.invoke(user_input)
    return response

# Gradio interface
iface = gr.Interface(
    fn=chat_with_llama,
    inputs=gr.Textbox(label="You"),
    outputs=gr.Textbox(label="LLaMA 3"),
    title="Chat with LLaMA 3",
    description="A simple chat interface using Ollama LLaMA 3"
)

iface.launch()
```

output:-

```
PROBLEMS 2 OUTPUT DEBUG CONSOLE TERMINAL PORTS
PS F:\Raj Singh> & C:/Users/User/AppData/Local/Programs/Python/Python313/python.exe "f:/Raj Singh/chat-gradio.py"
* Running on local URL: http://127.0.0.1:7860
* To create a public link, set `share=True` in `launch()`.
PS F:\Raj Singh> & C:/Users/User/AppData/Local/Programs/Python/Python313/python.exe "f:/Raj Singh/chat-gradio.py"
```

Open in browser :-

As shown below <http://127.0.0.1:7860/>

## Chat with LLaMA 3

A simple chat interface using Ollama LLaMA 3

You

how are you ko hindi me transalte it



Clear

Submit

LLaMA 3

I'd be happy to translate our conversation into Hindi for you.

Here's the translation:

मैं कैसे हूँ?

Translation: Main kaise hoon? (How am I?)

You can ask me any question, and I'll do my best to respond in Hindi. Keep model, so my responses may not be perfect, but I'll do my best to assist yo

What would you like to know or talk about in Hindi?

Flag