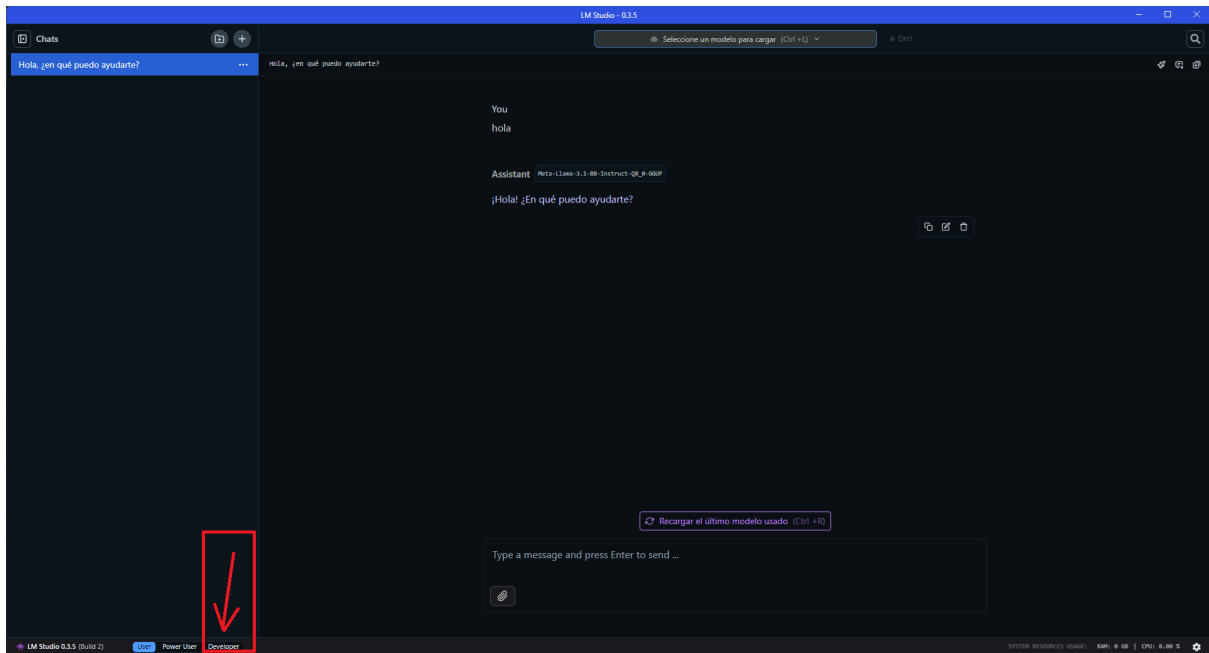
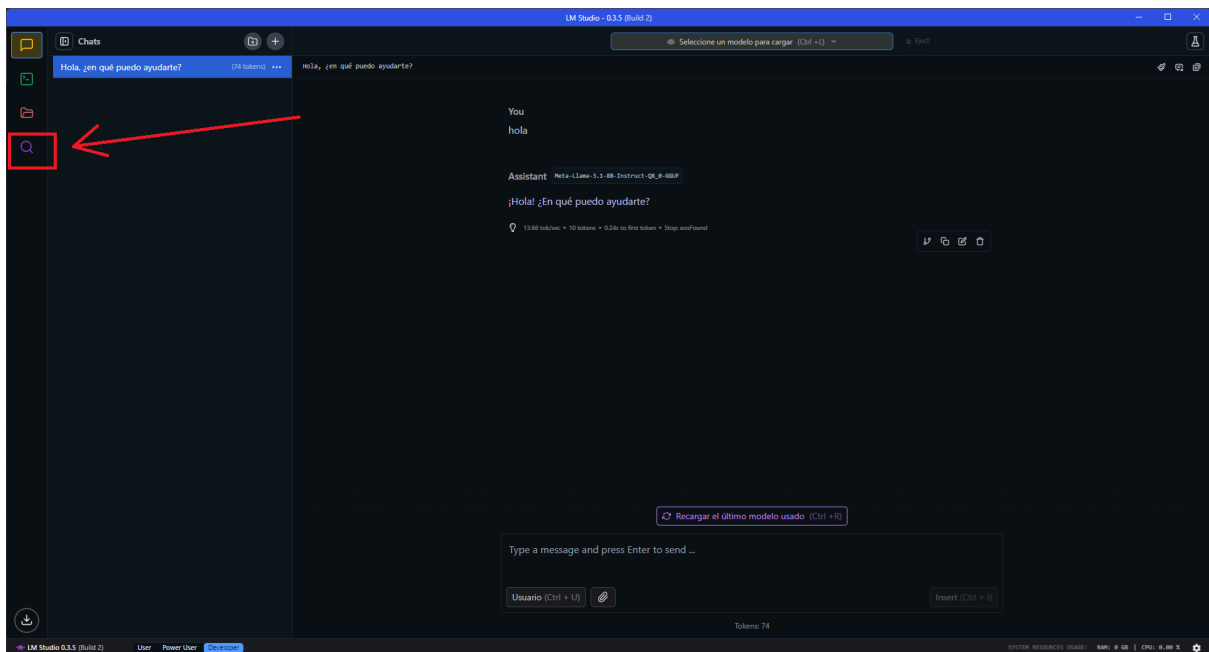


## Paso 1: modo “developer”:



## Paso 2: Buscamos un modelo para descargar.



### Paso 3: descargar un modelo

The screenshot shows the LM Studio Mission Control interface. The search bar at the top contains 'llama 3.1'. The search results list several models, with 'Meta Llama 3.1 8B' highlighted. The details panel on the right shows the model's architecture (Llama), parameters (8B), and status (222 likes, 41511 downloads). It also includes a 'Download' button with a size of 8.54 GB. Red arrows point to the search bar, the highlighted model, the 'Last updated' date, and the 'Download' button.

Model Search: Ctrl + M  
Runtimes: Ctrl + R

Showing 2817 models

Mejor Coincidencia

Meta Llama 3.1 8B

LM Studio Staff Pick

The latest in Meta's long-running Llama series, Llama 3.1 is another jack of all trades and master of some, now in 8 languages and up to 128k tokens.

Architecture: Llama  
Params: 8B  
Stats: 222 likes, 41511 downloads  
Last updated: 115 days ago

6 download options available

Q8\_0 Meta Llama 3.1 8B Instruct 8B 8.54 GB

Model Readme

Pulled from the model's repository

Community Model> Llama 3.1 8B Instruct by Meta

LM Studio Community models highlights program. Highlighting new & noteworthy models by the community. Join the conversation on Discord.

Model creator: meta-llama  
Original model: Meta-Llama-3.1-8B-Instruct  
GGUF quantization: provided by bartowski based on llama.cpp release b3472

Important: Requires LM Studio version 0.2.29, available now [here!](#)

Model Summary:

Llama 3.1 is an update to the previously released family of Llama 3 models. It is

Cancel Download 8.54 GB

En nuestro caso tenemos nuestras propia versiones:

The screenshot shows the LM Studio Mission Control interface. The search bar at the top contains 'Anrom97'. The search results list two models, with 'LLama-3.2-WIS-v1.0-GGUF' highlighted. The details panel on the right shows the model's repository (Anrom97/LLama-3.2-WIS-v1.0-GGUF), status (0 likes, 27 downloads), and last updated date (2 days ago). It also includes a 'Download' button with a size of 3.42 GB. The 'Model Readme' section provides information about the model's development, license, and training.

Model Search: Ctrl + M  
Runtimes: Ctrl + R

Showing 2 models

Mejor Coincidencia

LLama-3.2-WIS-v1.0-GGUF

Repository: Anrom97/LLama-3.2-WIS-v1.0-GGUF

Status: 0 likes, 27 downloads  
Last updated: 2 days ago

One download option available

Q8\_0 Unsloth 3.42 GB

Model Readme config.json

Pulled from the model's repository

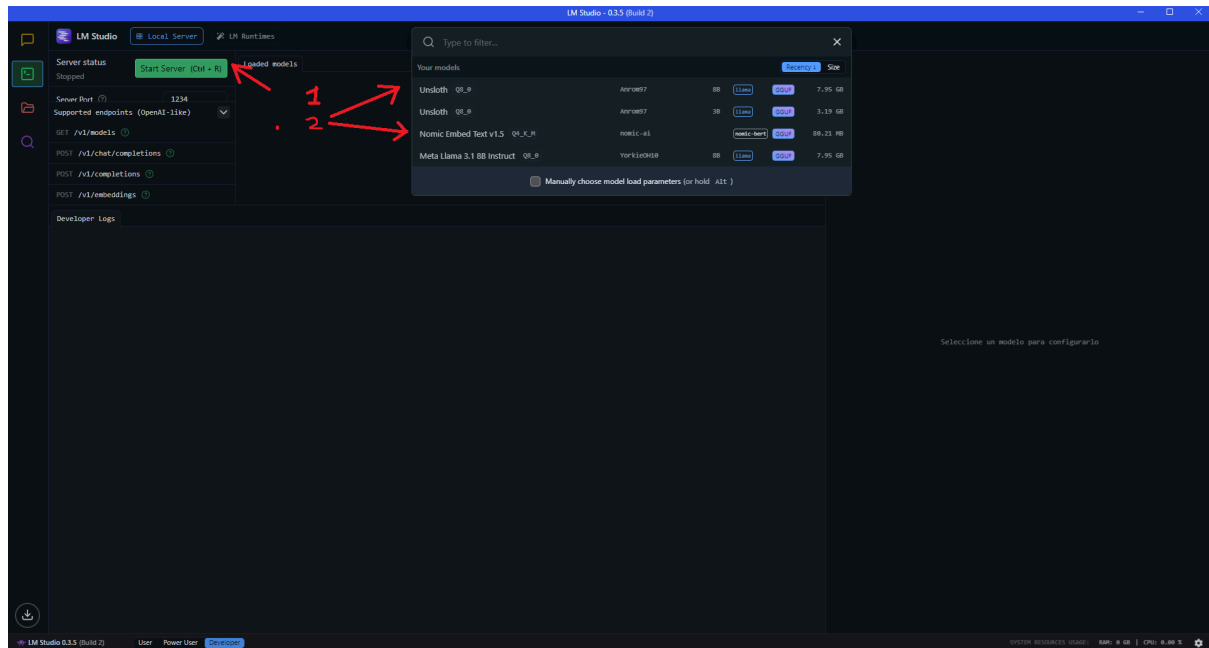
Uploaded model

- Developed by: Anrom97
- License: apache-2.0
- Finetuned from model: unsloth/llama-3.2-3b-instruct-bnb-4bit

This llama model was trained 2x faster with Unsloth and Huggingface's TRL library.

Cancel Use in New Chat

#### Paso 4: Levantar el servidor.



Primero: Cargamos el modelo (Llama 3.1 WIS).

Segundo: Cargamos el modelo de embedding (Nomic Embed Text 1.5)

Tercero: Oprimimos en "Start Server".

Correr el chatbot:

Librerías necesarias (instalar antes de correr los scripts):

```
pip install streamlit
```

```
pip install requests
```

```
pip install chromadb
```

```
pip install pypdf2
```

```
pip install python-docx
```

También se puede instalar todas las dependencias de una sola vez utilizando el archivo requirements.txt:

```
pip install -r requirements.txt
```

Crear la base de datos de vectores de RAG:

```
python create_database.py
```

Correr el chatbot (Es necesario tener el servidor de LLMStudio levantado):

```
streamlit run app.py
```