# Unit 5.2: Large Language Models

Utrecht Summer School: Programming with Python 2025

2025-07-11

Universiteit Utrecht

# Table of Contents

1. What is an LLM
2. Ethics
3. HuggingChat
4. Colab
5. Local LLMs
6. Conclusions

# Stupid is as stupid does

- What do you think is the next word that I will...?

# Stupid is as stupid does

- What do you think is the next word that I will WRITE?

## Stupid is as stupid does

- What do you think is the next word that I will WRITE?
- What do you think is the next word that I will SAY?

# Stupid is as stupid does

- What do you think is the next word that I will WRITE?
- What do you think is the next word that I will SAY?
- How did you guess?

# LLMs as coding agents

```
a = [2, 4, 5]
```

# LLMs as coding agents

```
a = [2, 4, 5]
# Compute the mean of the list
```

Universiteit Utrecht

# LLMs as coding agents

```
a = [2, 4, 5]
# Compute the mean of the list
```

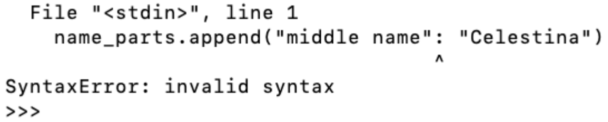
Can the LLM auto-complete?

# LLMs as bug fixers

```
# Create a dictionary with elements in the name of a person
name_parts = {"given name":  "Consolacion", "surname":  "Valenzuela"}
# Add a middle name
name_parts.append("middle name":  "Celestina")
```

# LLMs as bug fixers

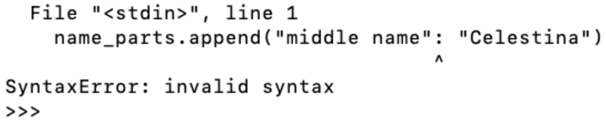
```
# Create a dictionary with elements in the name of a person
name_parts = {"given name": "Consolacion", "surname": "Valenzuela"}
# Add a middle name
name_parts.append("middle name": "Celestina")
```



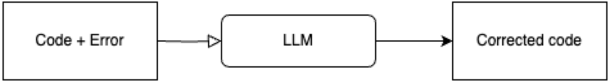
```
  File "<stdin>", line 1
    name_parts.append("middle name": "Celestina")
                                   ^
SyntaxError: invalid syntax
>>>
```

# LLMs as bug fixers

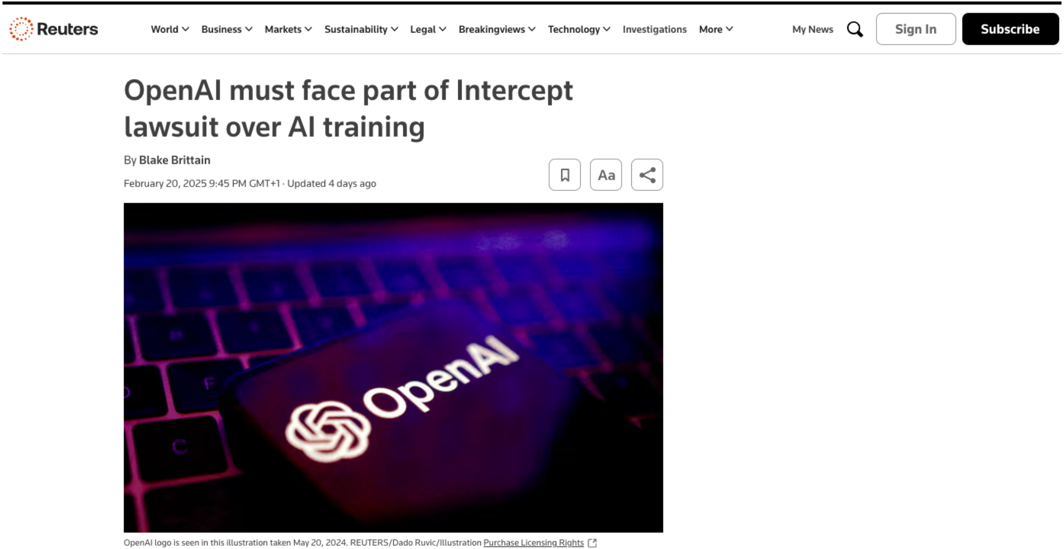
```
# Create a dictionary with elements in the name of a person
name_parts = {"given name": "Consolacion", "surname": "Valenzuela"}
# Add a middle name
name_parts.append("middle name": "Celestina")
```



```
  File "<stdin>", line 1
    name_parts.append("middle name": "Celestina")
                                   ^
SyntaxError: invalid syntax
>>>
```



# Training data



## OpenAI must face part of Intercept lawsuit over AI training

By Blake Brittain

February 20, 2025 9:45 PM GMT+1 · Updated 4 days ago

OpenAI logo is seen in this illustration taken May 20, 2024. REUTERS/Dado Ruvic/Illustration Purchase Licensing Rights

# Environment

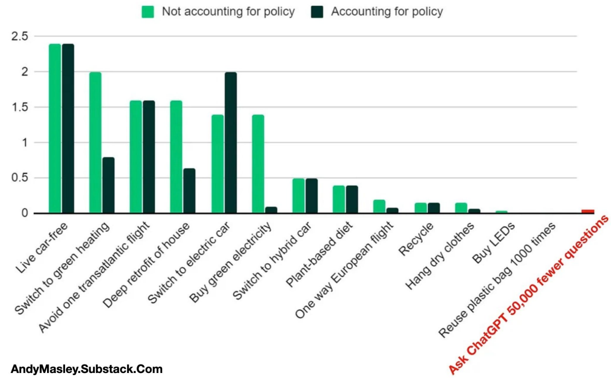


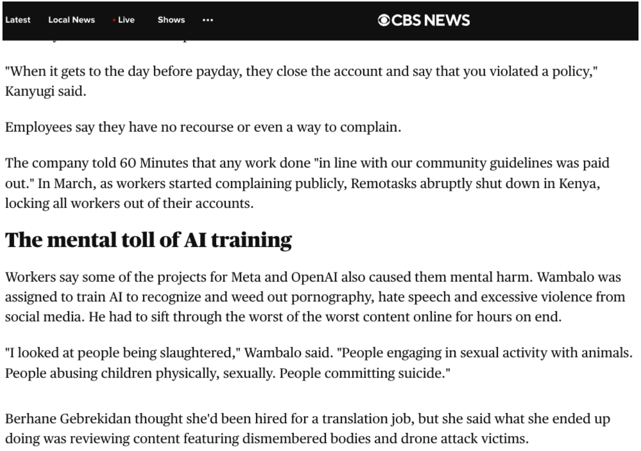
***Figure 4.*** Tonnes of $CO_2$ avoided by selected personal lifestyle decisions accounting for government policy

**AndyMasley.Substack.Com**
**Original climate graph taken from Founders Pledge**

[1]

[1] Source (January 2025): https://open.substack.com/pub/andymasley/p/individual-ai-use-is-not-bad-for

# Labor conditions



"When it gets to the day before payday, they close the account and say that you violated a policy," Kanyugi said.

Employees say they have no recourse or even a way to complain.

The company told 60 Minutes that any work done "in line with our community guidelines was paid out." In March, as workers started complaining publicly, Remotasks abruptly shut down in Kenya, locking all workers out of their accounts.

## The mental toll of AI training

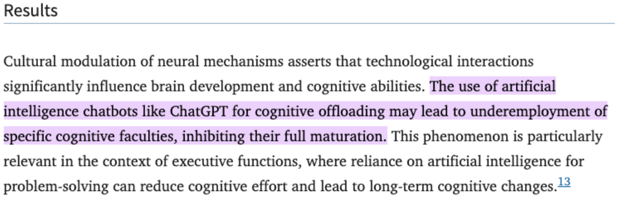
Workers say some of the projects for Meta and OpenAI also caused them mental harm. Wambalo was assigned to train AI to recognize and weed out pornography, hate speech and excessive violence from social media. He had to sift through the worst of the worst content online for hours on end.

"I looked at people being slaughtered," Wambalo said. "People engaging in sexual activity with animals. People abusing children physically, sexually. People committing suicide."

Berhane Gebrekidan thought she'd been hired for a translation job, but she said what she ended up doing was reviewing content featuring dismembered bodies and drone attack victims. [2]

[2] Source (November 2024): https://www.cbsnews.com/news/ai-work-kenya-exploitation-60-minutes/

# Effect on cognition



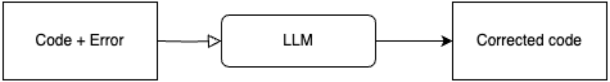
### Results

Cultural modulation of neural mechanisms asserts that technological interactions significantly influence brain development and cognitive abilities. The use of artificial intelligence chatbots like ChatGPT for cognitive offloading may lead to underemployment of specific cognitive faculties, inhibiting their full maturation. This phenomenon is particularly relevant in the context of executive functions, where reliance on artificial intelligence for problem-solving can reduce cognitive effort and lead to long-term cognitive changes.[13]

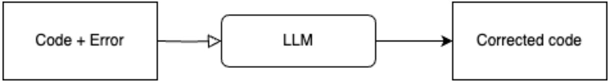
[3]

[3] Source (June 2024): Dubey et al. Redefining Cognitive Domains in the Era of ChatGPT: A Comprehensive Analysis of Artificial Intelligence's Influence and Future Implications. Med Res Arch. 2024 Jun;12(6):5383. doi: 10.18103/mra.v12i6.5383

# Uploading restricted data

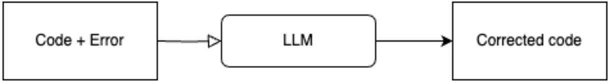


# Uploading restricted data



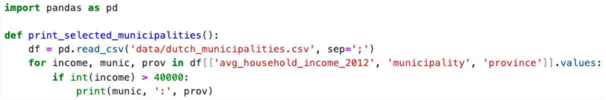
- What if error contains sensitive information?

# Uploading restricted data



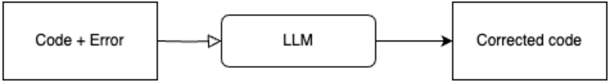
Code + Error → LLM → Corrected code

- What if error contains sensitive information?

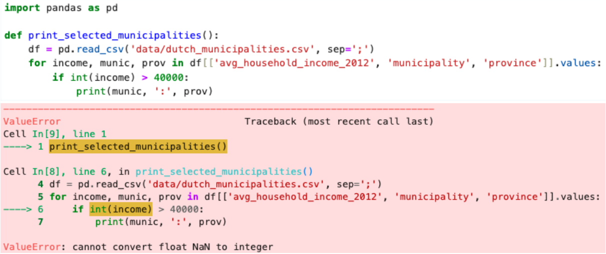


```
import pandas as pd

def print_selected_municipalities():
    df = pd.read_csv('data/dutch_municipalities.csv', sep=';')
    for income, munic, prov in df[['avg_household_income_2012', 'municipality', 'province']].values:
        if int(income) > 40000:
            print(munic, ':', prov)
```

# Uploading restricted data



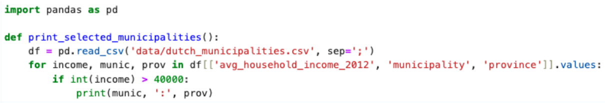
- What if error contains sensitive information?



```
import pandas as pd

def print_selected_municipalities():
    df = pd.read_csv('data/dutch_municipalities.csv', sep=';')
    for income, munic, prov in df[['avg_household_income_2012', 'municipality', 'province']].values:
        if int(income) > 40000:
            print(munic, ':', prov)
```

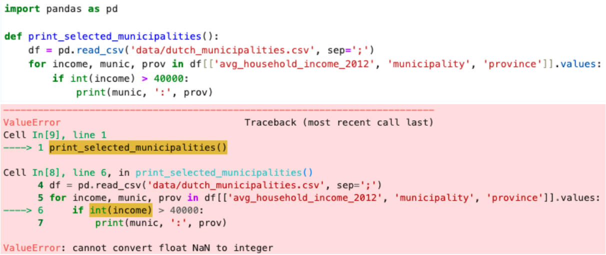
```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Cell In[9], line 1
----> 1 print_selected_municipalities()

Cell In[8], line 6, in print_selected_municipalities()
      4 df = pd.read_csv('data/dutch_municipalities.csv', sep=';')
      5 for income, munic, prov in df[['avg_household_income_2012', 'municipality', 'province']].values:
----> 6     if int(income) > 40000:
      7         print(munic, ':', prov)

ValueError: cannot convert float NaN to integer
```

# Uploading restricted data



```
import pandas as pd

def print_selected_municipalities():
    df = pd.read_csv('data/dutch_municipalities.csv', sep=';')
    for income, munic, prov in df[['avg_household_income_2012', 'municipality', 'province']].values:
        if int(income) > 40000:
            print(munic, ':', prov)
```

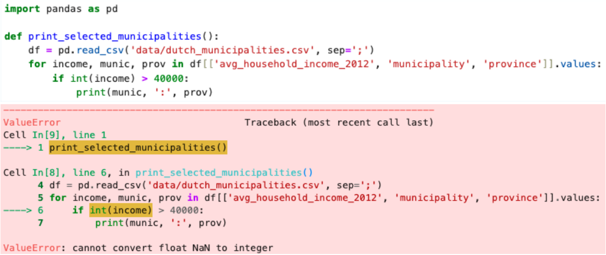
# Uploading restricted data



```
import pandas as pd

def print_selected_municipalities():
    df = pd.read_csv('data/dutch_municipalities.csv', sep=';')
    for income, munic, prov in df[['avg_household_income_2012', 'municipality', 'province']].values:
        if int(income) > 40000:
            print(munic, ':', prov)
```

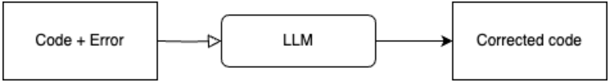
```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Cell In[9], line 1
----> 1 print_selected_municipalities()

Cell In[8], line 6, in print_selected_municipalities()
      4 df = pd.read_csv('data/dutch_municipalities.csv', sep=';')
      5 for income, munic, prov in df[['avg_household_income_2012', 'municipality', 'province']].values:
----> 6     if int(income) > 40000:
      7         print(munic, ':', prov)

ValueError: cannot convert float NaN to integer
```

# Uploading restricted data

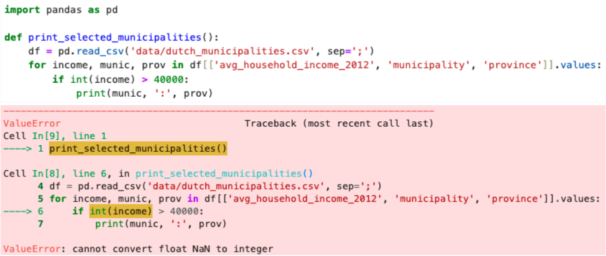


```
import pandas as pd

def print_selected_municipalities():
    df = pd.read_csv('data/dutch_municipalities.csv', sep=';')
    for income, munic, prov in df[['avg_household_income_2012', 'municipality', 'province']].values:
        if int(income) > 40000:
            print(munic, ':', prov)
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Cell In[9], line 1
----> 1 print_selected_municipalities()

Cell In[8], line 6, in print_selected_municipalities()
      4 df = pd.read_csv('data/dutch_municipalities.csv', sep=';')
      5 for income, munic, prov in df[['avg_household_income_2012', 'municipality', 'province']].values:
----> 6     if int(income) > 40000:
      7         print(munic, ':', prov)

ValueError: cannot convert float NaN to integer
```

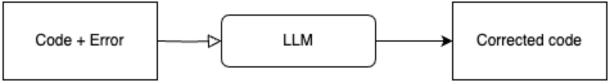


# Uploading restricted data



```
import pandas as pd

def print_selected_municipalities():
    df = pd.read_csv('data/dutch_municipalities.csv', sep=';')
    for income, munic, prov in df[['avg_household_income_2012', 'municipality', 'province']].values:
        if int(income) > 40000:
            print(munic, ':', prov)
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Cell In[9], line 1
----> 1 print_selected_municipalities()

Cell In[8], line 6, in print_selected_municipalities()
      4 df = pd.read_csv('data/dutch_municipalities.csv', sep=';')
      5 for income, munic, prov in df[['avg_household_income_2012', 'municipality', 'province']].values:
----> 6     if int(income) > 40000:
      7         print(munic, ':', prov)

ValueError: cannot convert float NaN to integer
```



Code + Error

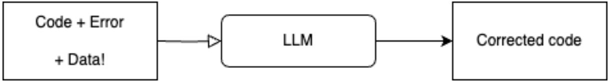
LLM

Corrected code

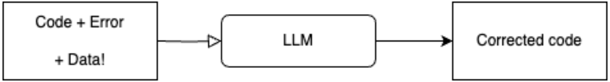
- Corrected code is useless / does not work

Universiteit Utrecht

# Uploading restricted data



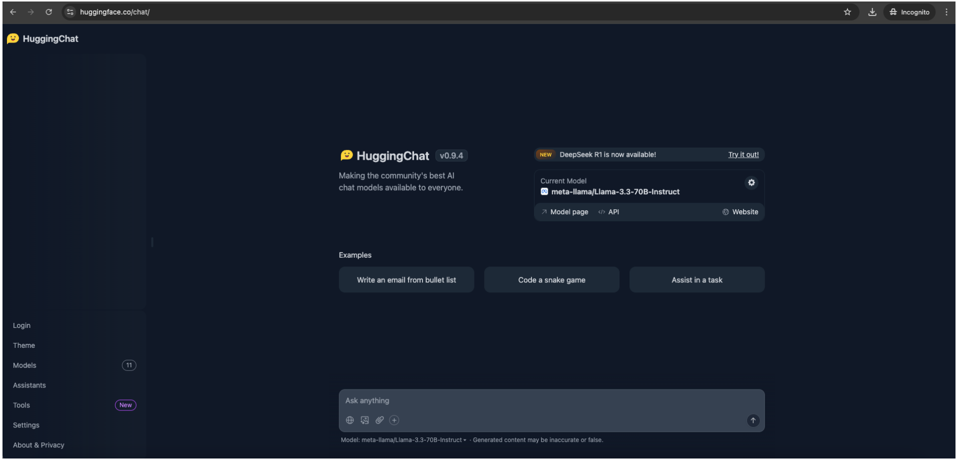
# Uploading restricted data



- May be illegal!

Universiteit Utrecht

# HuggingChat



HuggingChat v0.9.4

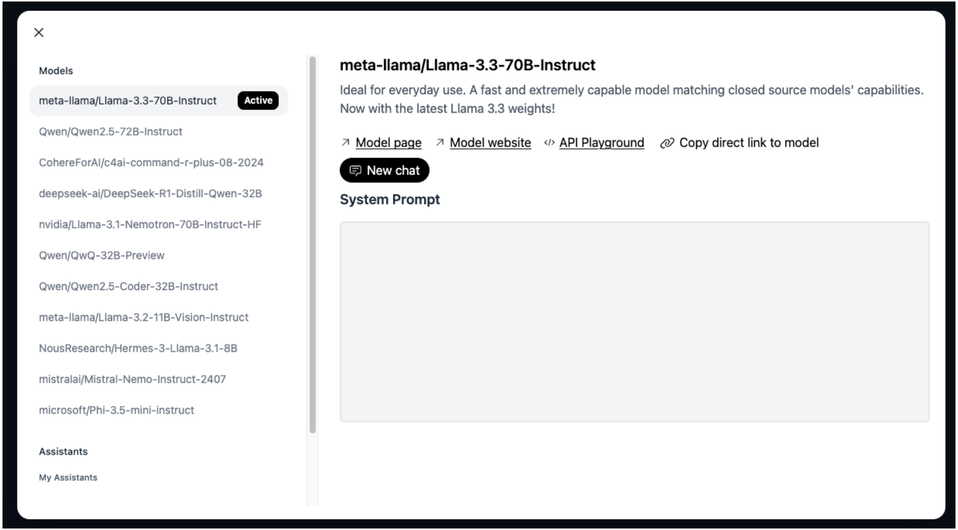
Making the community's best AI chat models available to everyone.

NEW DeepSeek R1 is now available! Try it out!

Current Model
meta-llama/Llama-3.3-70B-Instruct

Model: meta-llama/Llama-3.3-70B-Instruct · Generated content may be inaccurate or false.

# Switch models



Models

- meta-llama/Llama-3.3-70B-Instruct
- Qwen/Qwen2.5-72B-Instruct
- CohereForAI/c4ai-command-r-plus-08-2024
- deepseek-ai/DeepSeek-R1-Distill-Qwen-32B
- nvidia/Llama-3.1-Nemotron-70B-Instruct-HF
- Qwen/QwQ-32B-Preview
- Qwen/Qwen2.5-Coder-32B-Instruct
- meta-llama/Llama-3.2-11B-Vision-Instruct
- NousResearch/Hermes-3-Llama-3.1-8B
- mistralai/Mistral-Nemo-Instruct-2407
- microsoft/Phi-3.5-mini-instruct

Assistants

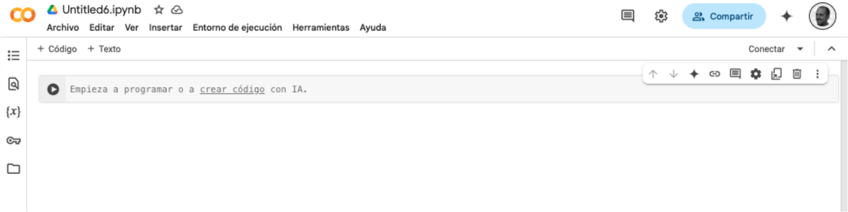
My Assistants

**meta-llama/Llama-3.3-70B-Instruct**

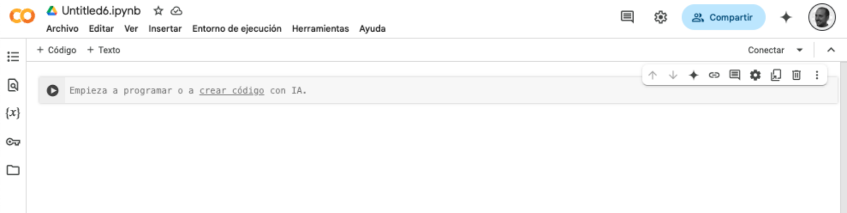
Ideal for everyday use. A fast and extremely capable model matching closed source models' capabilities. Now with the latest Llama 3.3 weights!

**System Prompt**

# Generative AI on Google Colab

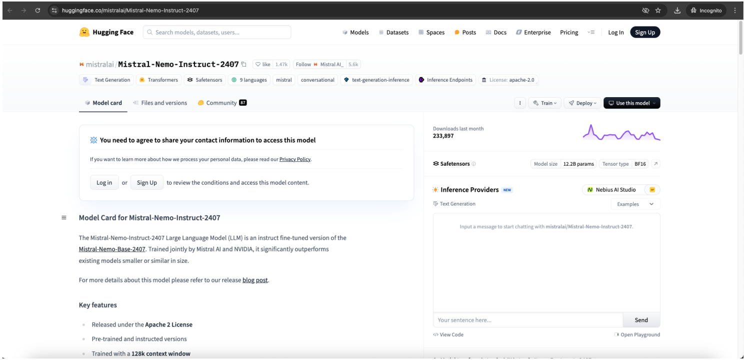


# Generative AI on Google Colab



Empieza a programar o a crear código con IA.

# Model cards



mistralai / **Mistral-Nemo-Instruct-2407**

**You need to agree to share your contact information to access this model**

If you want to learn more about how we process your personal data, please read our Privacy Policy.

Log in or Sign Up to review the conditions and access this model content.

Downloads last month
**233,897**

**Safetensors** Model size: 12.2B params · Tensor type: BF16

**Inference Providers** NEW · Nebius AI Studio

Text Generation

Input a message to start chatting with **mistralai/Mistral-Nemo-Instruct-2407**.

Your sentence here...

### Model Card for Mistral-Nemo-Instruct-2407

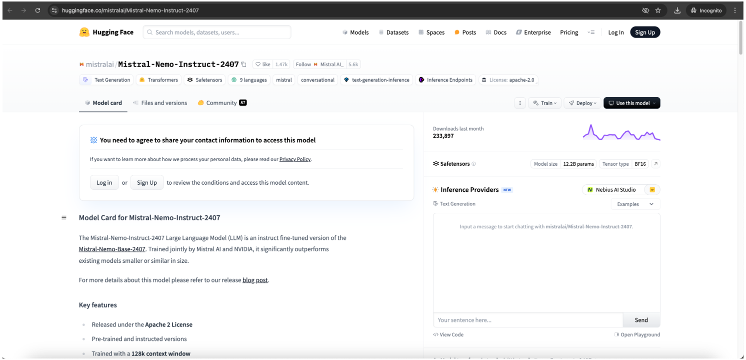
The Mistral-Nemo-Instruct-2407 Large Language Model (LLM) is an instruct fine-tuned version of the Mistral-Nemo-Base-2407. Trained jointly by Mistral AI and NVIDIA, it significantly outperforms existing models smaller or similar in size.

For more details about this model please refer to our release blog post.

#### Key features

- Released under the **Apache 2 License**
- Pre-trained and instructed versions
- Trained with a **128k context window**

# Model cards



**You need to agree to share your contact information to access this model**

If you want to learn more about how we process your personal data, please read our Privacy Policy.

Log in or Sign Up to review the conditions and access this model content.

Input a message to start chatting with mistralai/Mistral-Nemo-Instruct-2407.

### Model Card for Mistral-Nemo-Instruct-2407

The Mistral-Nemo-Instruct-2407 Large Language Model (LLM) is an instruct fine-tuned version of the Mistral-Nemo-Base-2407. Trained jointly by Mistral AI and NVIDIA, it significantly outperforms existing models smaller or similar in size.

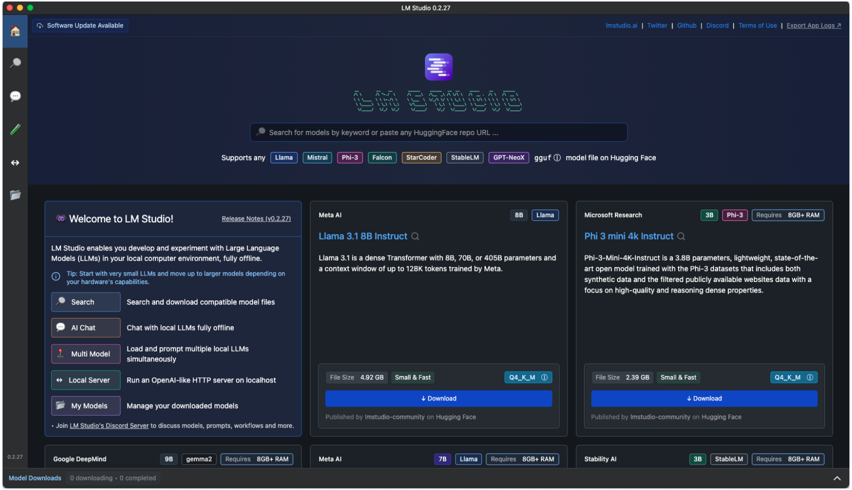
For more details about this model please refer to our release blog post.

#### Key features

- Released under the **Apache 2 License**
- Pre-trained and instructed versions
- Trained with a **128k context window**

No model card for Gemini!

# Using an LLM locally (no server)



LM Studio 0.2.27

Software Update Available
lmstudio.ai | Twitter | Github | Discord | Terms of Use | Export App Logs ↗
LM STUDIO
Search for models by keyword or paste any HuggingFace repo URL ...

Supports any Llama Mistral Phi-3 Falcon StarCoder StableLM GPT-NeoX gguf ⓘ model file on Hugging Face

**Welcome to LM Studio!** Release Notes (v0.2.27)

LM Studio enables you develop and experiment with Large Language Models (LLMs) in your local computer environment, fully offline.

ⓘ Tip: Start with very small LLMs and move up to larger models depending on your hardware's capabilities.

Search — Search and download compatible model files

AI Chat — Chat with local LLMs fully offline

Multi Model — Load and prompt multiple local LLMs simultaneously

Local Server — Run an OpenAI-like HTTP server on localhost

My Models — Manage your downloaded models

• Join LM Studio's Discord Server to discuss models, prompts, workflows and more.

Meta AI 8B Llama

**Llama 3.1 8B Instruct**

Llama 3.1 is a dense Transformer with 8B, 70B, or 405B parameters and a context window of up to 128K tokens trained by Meta.

File Size 4.92 GB Small & Fast Q4_K_M ⓘ

↓ Download

Published by lmstudio-community on Hugging Face

Microsoft Research 3B Phi-3 Requires 8GB+ RAM

**Phi 3 mini 4k Instruct**

Phi-3-Mini-4K-Instruct is a 3.8B parameters, lightweight, state-of-the-art open model trained with the Phi-3 datasets that includes both synthetic data and the filtered publicly available websites data with a focus on high-quality and reasoning dense properties.

File Size 2.39 GB Small & Fast Q4_K_M ⓘ

↓ Download

Published by lmstudio-community on Hugging Face

Google DeepMind 9B gemma2 Requires 8GB+ RAM

Meta AI 7B Llama Requires 8GB+ RAM

Stability AI 3B StableLM Requires 8GB+ RAM

# Time for a demo

But before we do that...

- Choose your poison carefully / Know what you choose

Universiteit Utrecht

# Time for a demo

But before we do that...

- Choose your poison carefully / Know what you choose
- Do not share private data

Universiteit Utrecht

# Time for a demo

But before we do that...

- Choose your poison carefully / Know what you choose
- Do not share private data
- Consider if you really need this