

Despliegue de modelos de IA

Alejandro Fernández Camello

Python Coruña

24 de febrero de 2024



1 Hugging Face

2 Modelos

3 Streamlit

4 Conclusiones



1 Hugging Face

2 Modelos

3 Streamlit

4 Conclusiones



¿Qué es Hugging Face?

- Empresa líder en inteligencia artificial y procesamiento del lenguaje natural.
- Modelo de negocio *open source*
- Ofrecen servicios *cloud* con sus librerías.
- Similar a GitHub, pero de modelos de inteligencia artificial.



Transformers

- Es la API principal de Hugging Face.
- Contiene modelos para procesar texto, imágenes, audio, etc.
- Puedes usar modelos muy avanzados en unas pocas líneas.



Ejemplo

```
from transformers import pipeline
```

```
classifier = pipeline("sentiment-analysis")
```

```
classifier("What's up?")
```



Índice

1 Hugging Face

2 Modelos

3 Streamlit

4 Conclusiones



Modelos

The screenshot shows the Hugging Face website's 'Models' page. The top navigation bar includes the Hugging Face logo, a search bar, and links for Models, Datasets, Spaces, Posts, Docs, Solutions, Pricing, Log In, and Sign Up. The left sidebar contains filters for Tasks (Multimodal, Computer Vision, Natural Language Processing), Libraries, Datasets, Languages, Licenses, and Other. The main content area displays a grid of model cards, each with a logo, name, description, update date, and performance metrics. The models are sorted by trending.

Model Name	Category	Update Date	Score	Popularity
stabilityai/stable-code-3b	Text Generation	Updated 7 days ago	± 4.77k	361
mistralai/Mixtral-8x7B-Instruct-v0.1	Text Generation	Updated Dec 15, 2023	± 1.09M	2.32k
h94/IP-Adapter-FaceID	Text-to-Image	Updated 6 days ago	± 187k	810
cagliostrolab/animagine-xl-3.0	Text-to-Image	Updated 8 days ago	± 81.2k	507
microsoft/phi-2	Text Generation	Updated 7 days ago	± 439k	2.52k
TencentARC/PhotoMaker	Text-to-Image	Updated 11 days ago	± 39k	186
InstantX/InstantID	Text-to-Image	Updated 3 days ago	± 11.7k	103
01-ai/yi-VL-34B	Text Generation	Updated about 4 hours ago	± 2.91k	98
stabilityai/stablelm-2-1.6b	Text Generation	Updated 1 day ago	± 2.13k	78
openai/whisper-large-v3	Automatic Speech Recognition	Updated 3 days ago	± 1.33M	1.49k
meta-llama/Llama-2-7b-chat-hf	Text Generation	Updated Nov 13, 2023	± 790k	2.51k
mistralai/Mixtral-8x7B-v0.1	Text Generation	Updated 4 days ago	± 134k	1.12k
NousResearch/Nous-Hermes-2-Mixtral-8x7B-DPO	Text Generation	Updated 8 days ago	± 7.42k	137
stabilityai/stable-diffusion-xl-base-1.0	Text-to-Image	Updated Oct 30, 2023	± 2.78M	4.2k
mlabonne/NeuralBeagle14-7B	Text Generation	Updated 2 days ago	± 2.95k	102
01-ai/yi-VL-6B	Text Generation	Updated about 4 hours ago	± 324	63

Figura: Captura de la pestaña para seleccionar los modelos



[Models](#)
[Datasets](#)
[Spaces](#)
[Posts](#)
[Docs](#)
[Solutions](#)
[Pricing](#)
[Log In](#)
[Sign Up](#)

Mixtral-8x7B-Instruct-v0.1
like 2.32k

[Text Generation](#)
[Transformers](#)
[Safetensors](#)
[5 languages](#)
[mixtral](#)
[text-generation-inference](#)
License: apache-2.0

[Model card](#)
[Files and versions](#)
[Community 110](#)

[Train](#)
[Deploy](#)
[Use in Transformers](#)

Model Card for Mixtral-8x7B

The Mixtral-8x7B Large Language Model (LLM) is a pretrained generative Sparse Mixture of Experts. The Mixtral-8x7B outperforms Llama 2 70B on most benchmarks we tested.

For full details of this model please read our [release blog post](#).

Warning

This repo contains weights that are compatible with [vLLM](#) serving of the model as well as Hugging Face [transformers](#) library. It is based on the original Mixtral [torrent release](#), but the file format and parameter names are different. Please note that model cannot (yet) be instantiated with HF.

Instruction format

This format must be strictly respected, otherwise the model will generate sub-optimal outputs.

The template used to build a prompt for the instruct model is defined as follows:

```
<s> [INST] Instruction [/INST] Model answer</s> [INST] Follow-up instruction [/IN
```

Downloads last month

1,085,380

[Safetensors](#)
[Model size 46.7B params](#)
[Tensor type BF16](#)

[Text Generation](#)

Inference API has been turned off for this model.

Spaces using mistralai/Mixtral-8x7B-Instruct-v0.1 319

ehristoforu/mixtral-46.7b-chat

Tomonlai/Mixtral-Chat

gospacedev/emoji-translator

HeshamHarooun/Arabic_mixtral_8-7b

pragneshbarik/ikigai-chat

Skier8402/mixtral-PDF-chat

smart-fellah/Smart-fellah

Cenaashoori/Mixtral-Chat

TogetherAI/TogetherChat-Coder

muhammedturan/TR-Chat-v1

meraGPT/meraKB

einfachall/Einfach.Chat

eskayML/Nurse_Nkuru

Statcal/STC-LLM

AFischer1985/AI-Interface

TheBlueberry-AI/Mist-Chat

4emes johann22/mixtral-test-46.7b-chat

broadfield/Mixtral-Agent

rishrai/mixtral

medmac01/smart_fellah

+ 299 Spaces

Figura: Captura de la pestaña de ejemplo de un modelo

Alejandro Fernández (Python Coruña)

Despliegue de modelos de IA

24 de febrero de 2024

9 / 15

Índice

1 Hugging Face

2 Modelos

3 Streamlit

4 Conclusiones



¿Qué es Streamlit?

- Streamlit: Framework de código abierto para crear aplicaciones web en Python.
- Fácil de usar: Permite crear interfaces de usuario interactivas con pocas líneas de código.
- Rápido para prototipar: Ideal para visualización de datos y construcción rápida de dashboards.
- Integración de datos y ML: Soporta diversas librerías de datos y *machine learning*.



Ejemplo

Código

```
import streamlit as st

st.title('Ejemplo con Streamlit')
st.write('Selecciona tus preferencias:')

color = st.selectbox(
    'Elige tu color favorito',
    ['Rojo', 'Azul', 'Verde']
)
st.write(f'Has elegido el color: {color}')

numero = st.slider('Selecciona un número', 1, 10)
st.write('Número seleccionado:', numero)

if st.button('Haz click aquí'):
    st.write('¡Botón presionado!')
```

Ejemplo con Streamlit

Selecciona tus preferencias:

Elige tu color favorito

Rojo

Has elegido el color: Rojo

Selecciona un número

1

1

10

Número seleccionado: 1

Haz click aquí

Figura: Interfaz de Streamlit



Índice

1 Hugging Face

2 Modelos

3 Streamlit

4 Conclusiones



Conclusiones

- Usando Hugging Face podemos integrar en nuestra aplicación modelos avanzados rápidamente.
- Con Streamlit creamos interfaces de usuario muy rápido en Python.
- Combinados nos permiten crear aplicaciones de inteligencia artificial en minutos.



¡Muchas gracias por haberme escuchado!

Tenéis disponible la presentación y el código en el siguiente enlace:

<https://github.com/alexfdex1010/talk-huggingface>

