



CUADERNOS DE CÓDIGO DE PYTHON PARA DOCENTES UNIVERSITARIOS

Sesión 2: SCOPUS & LLM

Jesús Antonio Alvarado Huayhuaz
DIRECCIÓN UNIVERSITARIA DE PERSONAL DOCENTE

Agosto-2024



Antes de iniciar...

Test # 1:
4 preguntas para
repasar la primera
sesión



PRE-ÁMBULO



coursera



Buscar en el curso

Buscar

Español



Ocultar menú

Introduction to Llama 2: A High Quality Open Source Large Language Model

✓ **Video:** Welcome to the
Class: Llama for Python
Programmers
3 min

✓ **Lectura:** Course
Syllabus
5 min

✓ **Cuadro de aviso de la
discusión:** Meet Your
Fellow Learners
10 min

✓ **Lectura:** Help Us Learn
About You
5 min

✓ **Video:** What is Llama 2?
An Open Source Large
Language Model (LLM)
8 min

laboratorio: Llama 2
Workspace
1 min

✓ **Video:** Llamas Llamas
Everywhere! Inference.

Módulo1 > What is Llama 2? An Open Source Large Language Model (LLM)

< Anterior Siguiente >

Generative AI

- It builds on top of these and other techniques from machine learning.
- It generates new output data in response to both an input data stream and a pre-trained corpus.

data stream and a pre-trained corpus.

M

1:07 / 8:57

What is Llama 2? An Open Source Large Language Model (LLM)

coach



J **Jesus**
Que es llama2?

Coach
Llama2 es una familia de modelos de lenguaje de código abierto y alta calidad. Estos modelos son parte de la inteligencia artificial generativa y se utilizan para generar texto a partir de una entrada de texto. Llama2 ha sido creado por Meta y se ha hecho disponible de forma gratuita para su descarga y uso en investigación y también para fines comerciales. Estos modelos son especialmente interesantes debido a su alta calidad y a la cantidad de parámetros que tienen, que se miden en miles de millones. Además, existen

Pregúntame lo que quieras

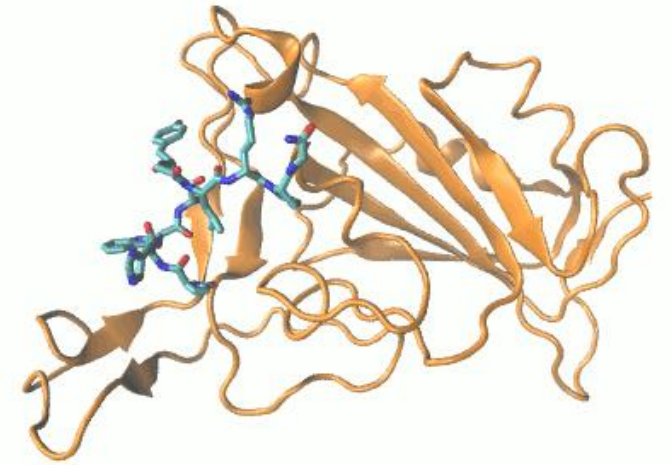


Coach está impulsado por IA, así que verifica si hay errores y no compartas información confidencial. Tus datos serán utilizados de acuerdo con el [Aviso de privacidad de Coursera](#).



Objetivos

- 1) Realizar consultas a SCOPUS desde el cuaderno de código para automatizar la búsqueda bibliográfica.
- 2) Procesamiento de información de repositorios académicos.
- 3) Aplicación de LLM para la consulta de artículos científicos.





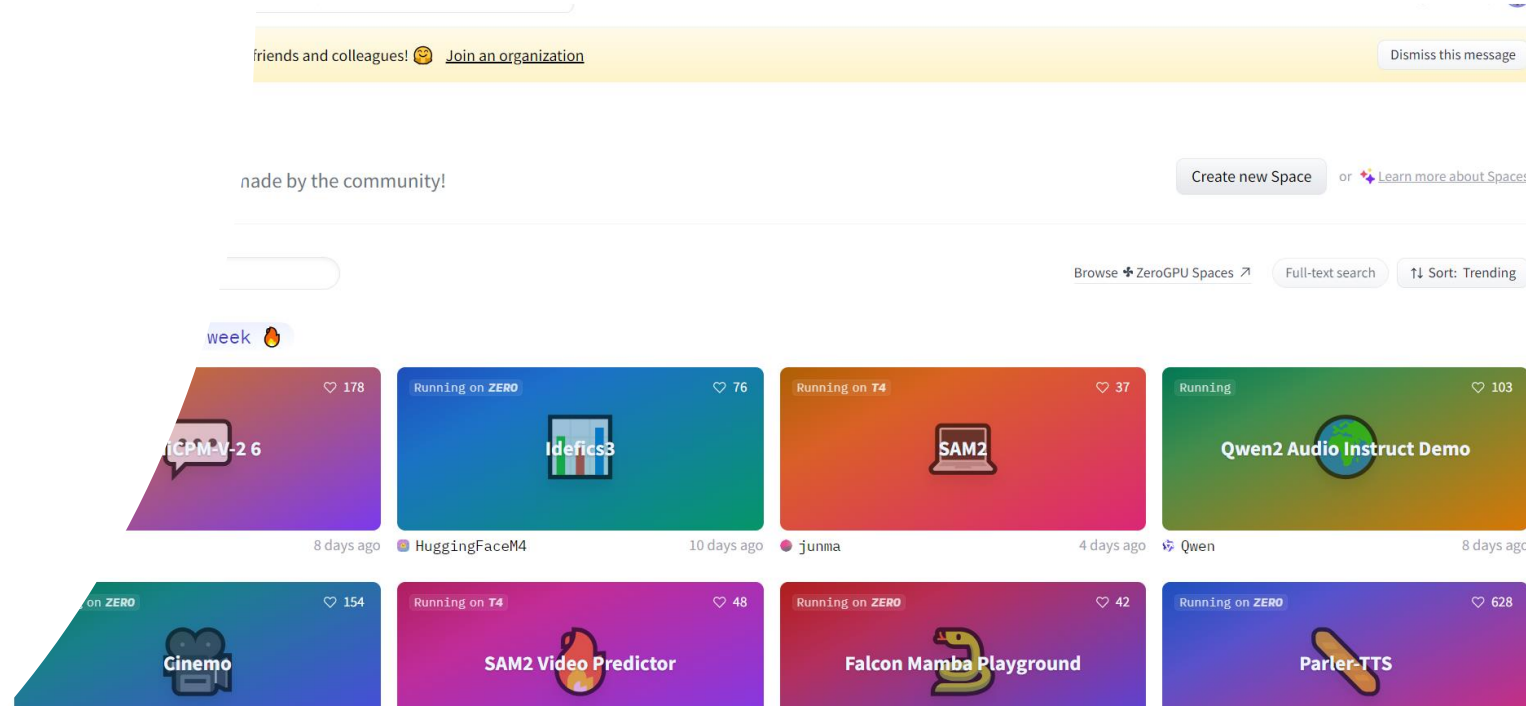
(Open Source) Large Language Model

- LLM es el acrónimo de "Large Language Model" (Modelo de Lenguaje Grande)
- LLM son modelos de lenguaje pre-entrenados que se utilizan para generar texto a partir de una entrada de texto
- Son parte de la inteligencia artificial generativa y se basan en técnicas de procesamiento de lenguaje natural y aprendizaje profundo



¿Qué aplicaciones te interesan explorar?

- Redacción científica
- Elaboración de clases
- Revisión de texto científico
- Ilustración de contenido científico
- Prácticas dirigidas



CARACTERÍSTICAS



GPT-4o, Claude 3.5, Mistral Large 2 y Llama 3.1.

1. Capacidades omnimodales: texto, voz, visión.
2. Respuestas ultrarrápidas: 320 milisegundos en promedio.
3. Dominio multilingüe: más de 50 idiomas.
4. Operaciones numéricas.
5. Métricas
 - MMLU: 88,7 (“sabe más que el graduado universitario promedio”)
 - HumanEval: 90,2 (codifica mejor que la mayoría de los humanos)



CARACTERÍSTICAS



GPT-4o, Claude 3.5, Mistral Large 2 y Llama 3.1.

- 1. Claude 3.5. Es el último proyecto publicado de Anthropic.
- 2. Métricas:
GPQA : Estableciendo nuevos récords en razonamiento a nivel de posgrado.
HumanEval : 91% de resolución de problemas.



	Claude 3.5 Sonnet	Claude 3 Opus	GPT-4o	Gemini 1.5 Pro	Llama-400b (early snapshot)
Graduate level reasoning <i>GPQA, Diamond</i>	59.4%* 0-shot CoT	50.4% 0-shot CoT	53.6% 0-shot CoT	—	—
Undergraduate level knowledge <i>MMLU</i>	88.7%** 5-shot	86.8% 5-shot	—	85.9% 5-shot	86.1% 5-shot
	88.3% 0-shot CoT	85.7% 0-shot CoT	88.7% 0-shot CoT	—	—
Code <i>HumanEval</i>	92.0% 0-shot	84.9% 0-shot	90.2% 0-shot	84.1% 0-shot	84.1% 0-shot
Multilingual math <i>MGSM</i>	91.6% 0-shot CoT	90.7% 0-shot CoT	90.5% 0-shot CoT	87.5% 8-shot	—

CARACTERÍSTICAS



GPT-4o, Claude 3.5, **Mistral Large 2** y Llama 3.1.

1. ¿Por qué está llamando la atención?

Ventana de contexto del token 32K

Maravilla multilingüe: habla inglés, francés, español, alemán e italiano con fluidez.

2. Métricas:

MMLU : Supera a algunos grandes nombres en tareas multilingües.

Programación y matemáticas: las mejores puntuaciones en HumanEval y GSM8K.



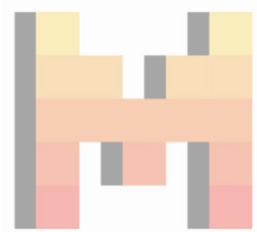
 Meta AI

CARACTERÍSTICAS



GPT-4o, Claude 3.5, Mistral Large 2 y **Llama 3.1**.

- Contexto - tokens 128K: Podría manejar toda una tesis.
- Disponibilidad en AWS, Databricks, Google Cloud, Hugging Face, Kaggle, IBM WatsonX, Microsoft Azure, NVIDIA NIM y Snowflake, y con soporte de plataformas de hardware ofrecidas por AMD, AWS, Dell, Intel, NVIDIA y Qualcomm.



Meta Llama 3 Instruct model performance

	Meta Llama 3 8B	Gemma 7B - It Measured	Mistral 7B Instruct Measured
MMLU 5-shot	68.4	53.3	58.4
GPQA 0-shot	34.2	21.4	26.3
HumanEval 0-shot	62.2	30.5	36.6
GSM-8K 8-shot, CoT	79.6	30.6	39.9
MATH 4-shot, CoT	30.0	12.2	11.0

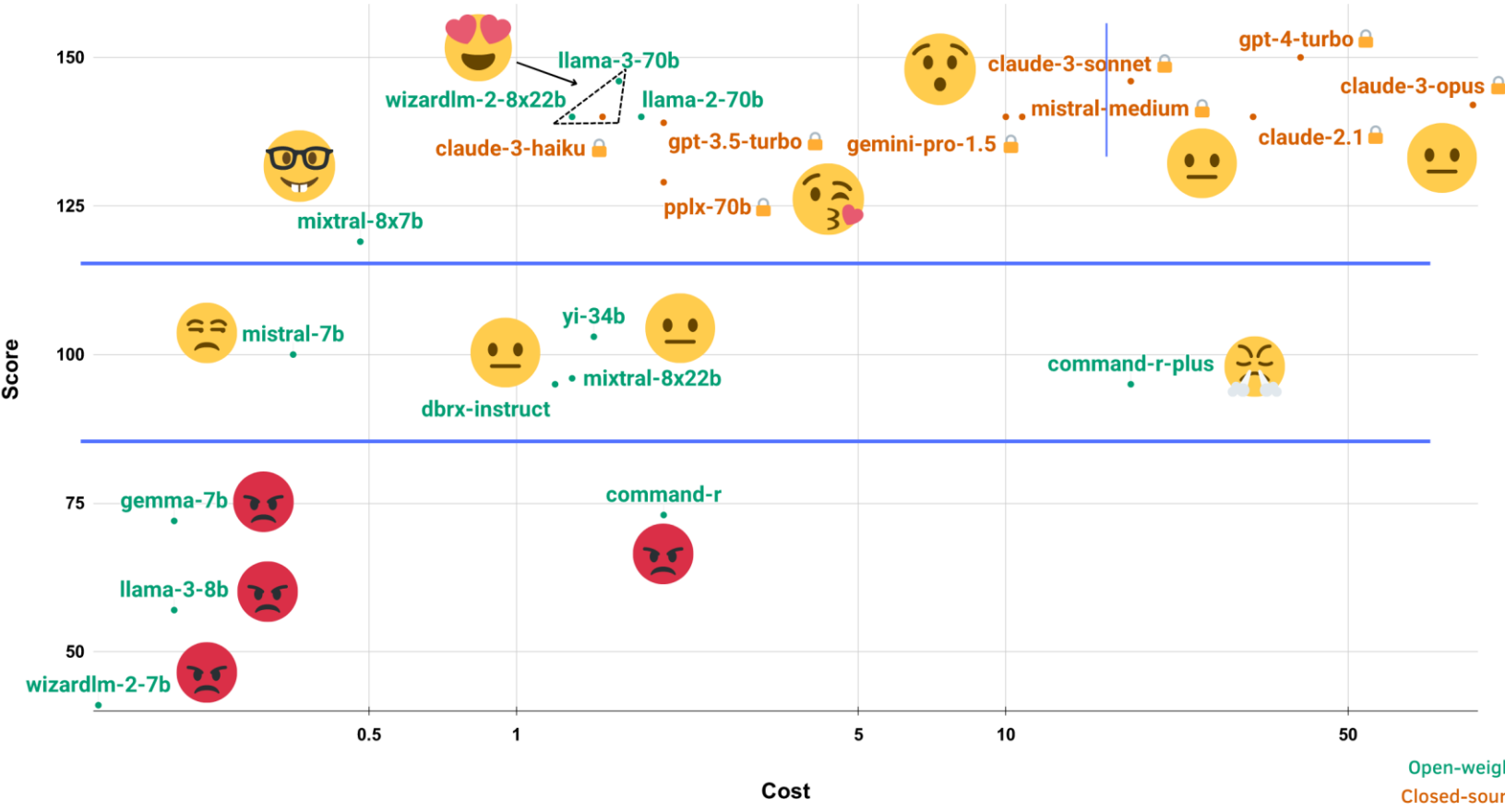
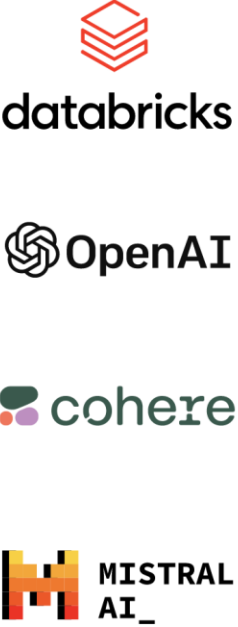
	Meta Llama 3 70B	Gemini Pro 1.5 Published	Claude 3 Sonnet Published
MMLU 5-shot	82.0	81.9	79.0
GPQA 0-shot	39.5	41.5 CoT	38.5 CoT
HumanEval 0-shot	81.7	71.9	73.0
GSM-8K 8-shot, CoT	93.0	91.7 11-shot	92.3 0-shot
MATH 4-shot, CoT	50.4	58.5 Minerva prompt	40.5



COMPARACIÓN DEL RENDIMIENTO



Which LLM among 138 can write tests for Java and Go?





Sesión práctica

Gracias

Dirección Universitaria de Personal Docente.
Vicerrectorado Académico.

2024



UNIVERSIDAD PERUANA
CAYETANO HEREDIA