

PROMPTING EN INTELIGENCIA ARTIFICIAL GENERATIVA

Módulo 2: Modelos de Lenguaje Amplio (LLMs - Large Language Models)

Instructor: Daniel M. Casas

Edición: Empresas

[Danimcasas.com](https://danimcasas.com) y Nironet formacio.

Máster en Inteligencia Artificial y Formador Tecnológico

danimcasas@inextalent.com

Actualizado: 2025



Módulo 2: Modelos de Lenguaje Amplio (LLMs)

1. ¿Qué es un LLM?

Un **Modelo de Lenguaje Amplio (Large Language Model, LLM)** es un sistema de inteligencia artificial entrenado con cantidades masivas de texto para **comprender y generar lenguaje humano** de manera coherente y contextualizada.

- Se basan en arquitecturas de **transformers**, un avance clave introducido por Google en 2017.
- Son capaces de realizar tareas complejas como la redacción, traducción, análisis de texto, generación de código o el razonamiento básico.
- Su principal característica es que **aprenden patrones complejos del lenguaje**, permitiendo no solo predecir la siguiente palabra, sino sostener conversaciones y crear documentos extensos.

Cómo Aprenden los LLMs

El desarrollo de un LLM se realiza en varias fases clave:

1. **Entrenamiento Inicial (*Pretraining*):**
 - Se alimenta al modelo con ingentes cantidades de texto de Internet, libros y artículos.
 - El objetivo es que el modelo aprenda la **probabilidad de aparición de palabras** en un contexto determinado.
2. **Ajuste Fino (*Fine-tuning*):**
 - Se entrena al modelo en datos más específicos o de nicho (por ejemplo, documentos legales o lenguaje administrativo) para refinar su estilo y precisión.
3. **RLHF (*Reinforcement Learning with Human Feedback*):**
 - Se ajusta el modelo mediante **retroalimentación humana** para asegurar que las respuestas sean útiles, seguras y que estén alineadas con valores éticos y sociales.

Ejemplos de LLMs Actuales

Modelo	Desarrollador	Enfoque Principal
GPT	OpenAI	El estándar más popular en aplicaciones de texto y chat conversacional.
Claude	Anthropic	Destacado por su seguridad, robustez y estilo más "razonado" en sus respuestas.
Gemini	Google	Enfoque nativamente multimodal (capaz de procesar texto, imagen y vídeo).
Llama 3	Meta	Modelo de código abierto muy extendido en proyectos de investigación y desarrollo.
Mistral AI	Mistral AI	Modelos conocidos por ser ligeros y eficientes, ideales para entornos profesionales.

Limitaciones de los LLMs

A pesar de su potencia, los LLMs presentan riesgos y limitaciones que deben conocerse:

- **Sesgos:** Heredan y pueden amplificar prejuicios sociales presentes en los datos de entrenamiento.
- **Alucinaciones:** Tienen la capacidad de **inventar información** que suena totalmente convincente, pero que es objetivamente falsa.
- **Consumo de Recursos:** El entrenamiento y la ejecución de estos modelos requieren una enorme cantidad de energía y una gran potencia computacional.
- **Seguridad y Privacidad:** Existe un riesgo inherente al exponer datos sensibles si se usan modelos no alojados de forma segura o sin las debidas precauciones.

2. Ejemplos de Prompting Aplicados

Aquí se muestran ejemplos de *prompts* diseñados para explorar las capacidades de los LLMs:

Ejemplo	Objetivo del Prompt	Prompt Sugerido
Ejemplo 1	Comparación de estilo	<i>Redacta una explicación clara y formal sobre qué es un acta administrativa.</i> (Recomendación: Ejecutar este prompt en dos modelos diferentes y comparar diferencias de estilo y claridad.)
Ejemplo 2	Razonamiento con contexto	<i>Actúa como un profesor universitario. Explica en 5 puntos las ventajas y los riesgos que presentan los modelos de lenguaje en el contexto de la administración pública.</i>
Ejemplo 3	Traducción administrativa	<i>Traduce el siguiente texto de castellano a gallego manteniendo un tono administrativo formal: [Pegar texto del documento a traducir].</i>

3. Resumen del Módulo

- Los **LLMs** son sistemas entrenados con volúmenes masivos de texto que utilizan la arquitectura **Transformers** para generar lenguaje humano.
- Su aprendizaje se basa en un proceso que incluye el *pretraining* (entrenamiento inicial), el *fine-tuning* (ajuste específico) y el **RLHF** (refuerzo con *feedback* humano).
- Los modelos más relevantes son **GPT, Claude, Gemini, Llama 3 y Mistral**.
- El uso en el sector público debe ser **crítico y responsable**, ya que estos modelos tienen limitaciones importantes como los **sesgos** y las **alucinaciones**.

4. Práctica (Entregable)

Actividad Módulo 2 – Análisis Comparativo de LLMs

El objetivo de esta práctica es evaluar la calidad y el estilo de respuesta de dos modelos de lenguaje diferentes para una tarea administrativa.

Pasos a seguir:

1. **Escoge dos LLMs diferentes** (ejemplo: ChatGPT y Claude, o ChatGPT y Gemini).
2. **Lanza el mismo *prompt*** en ambos modelos:
 - “Redacta un comunicado oficial sobre la implantación de un nuevo servicio digital en un ayuntamiento.”
3. **Copia los dos resultados** en un documento Word/PDF.
4. **Realiza un análisis comparativo** basándote en las siguientes preguntas:
 - ¿Cuál de los dos comunicados es más claro en su mensaje?
 - ¿Cuál tiene un mejor tono administrativo (formal, objetivo, preciso)?
 - ¿Detectas algún error de coherencia o alguna exageración (*alucinación*) en las respuestas?
5. **Conclusión:** ¿Qué modelo utilizarías en tu trabajo diario para este tipo de tarea y por qué?

Formato de entrega: Documento en Word o PDF subido a la plataforma Moodle.

5. Fuentes Recomendadas (Vídeos y Lecturas Complementarias)

Fuente / institución	Título / Descripción	URL
DeepLearning.AI + OpenAI	ChatGPT Prompt Engineering for Developers (curso corto)	https://www.youtube.com/watch?v=tRvcAdqsJWo
Stanford (CS224N 2023)	Lecture 10 – Prompting & Reinforcement Learning from Human Feedback	https://www.youtube.com/watch?v=SXpJ9EmG3s4
MIT (6.S087, 2024)	Foundation Models & Generative AI (playlist oficial del curso)	https://www.youtube.com/playlist?list=PLXV9Vh2jYcjbv67sXNDJiO8MWLA3ZJKR
MIT Sloan	Primer on ChatGPT (incluye mejores prácticas de prompting)	https://www.youtube.com/watch?v=4fThho0NvA0
Harvard (HILT)	Harvard Generative AI Tools: Demo and Discussion (sección sobre prompting/temperature)	https://www.youtube.com/watch?v=61zn8Q6IK08

PROMPTING EN INTELIGENCIA ARTIFICIAL GENERATIVA - Empresas
Instructor: Daniel M. Casas - actualizado 2025

Harvard (CS50x 2025)	Artificial Intelligence (incluye sección de Prompt Engineering)	https://www.youtube.com/watch?v=1HuD2rYeOLg
DeepLearning.AI + OpenAI	Building Systems with the ChatGPT API (aplica prompting en sistemas)	https://www.youtube.com/watch?v=BunESRhYhec
MIT IAP / AI ML Club	Next-Generation Prompt Engineering (demo + discusión)	https://www.youtube.com/watch?v=ML5zDg5SKzY
MIT (6.5940, 2023)	Efficient Fine-tuning and Prompt Engineering	https://www.youtube.com/watch?v=vOPwwRCZ8q8