Programación con IA Generativa

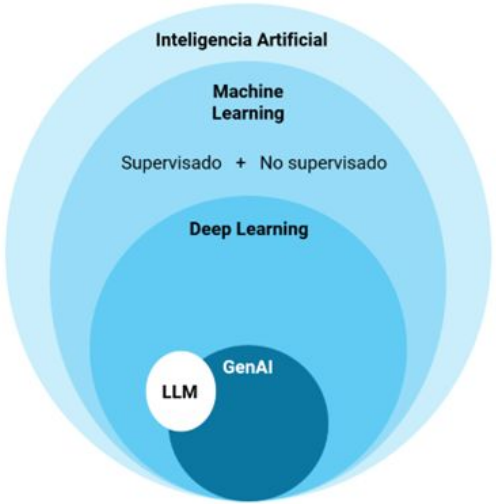
Modelos LLM para asistencia de código

Índice

1 Modelos LLM

IA Generativa

- Subcampo de la **IA**
- Permite generar **contenido** original e innovador
- Imágenes, texto, música o **código**




OpenAI


Claude
BY
ANTHROPIC


Grok


Gemini
Google

Tareas de programación de uso general

Usa estos modelos para **tareas de desarrollo comunes** que requieren un equilibrio de calidad, velocidad y rentabilidad.

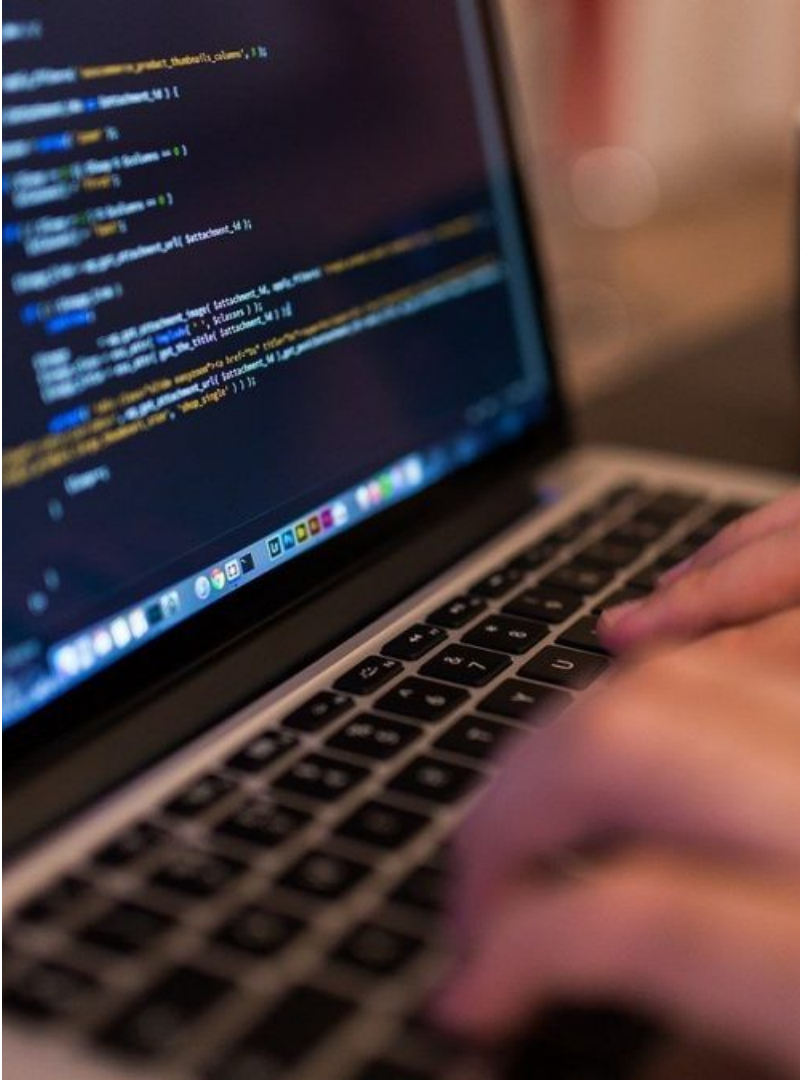
Estos modelos son un buen valor predeterminado **cuando no tienes requisitos específicos.**

Tareas de programación de uso general

Usar uno de estos modelos cuando queremos:

- Escribir o revisar funciones, archivos cortos o diferencias de código.
- Generar documentación, comentarios o resúmenes.
- Explicar rápidamente errores o un comportamiento inesperado.
- Trabajar en un entorno de programación que no sea en inglés

Modelo	Por qué es una buena opción
GPT-4.1	Valor predeterminado confiable para la mayoría de las tareas de programación y escritura. Rápido, preciso y funciona bien entre lenguajes y marcos.
GPT-4o	Rápido, menos preciso que GPT-4.1.
Claude Sonnet 3.7	Genera una salida clara y estructurada. Sigue instrucciones de formato y mantiene un estilo coherente.
Gemini 2.0 Flash	Rápido y rentable. Adecuado para preguntas rápidas, fragmentos de código cortos y tareas de escritura ligeras.
o4-mini	Optimizado para la velocidad y la rentabilidad. Ideal para sugerencias en tiempo real con una sobrecarga de uso baja.



Tareas de ayuda rápida, tareas sencillas o repetitivas

Estos modelos están **optimizados para la velocidad** y la capacidad de respuesta.

Son ideales para ediciones rápidas, funciones de utilidad, ayuda de sintaxis y creación de prototipos ligeros.

Obtendrás respuestas rápidas sin tener que esperar por cadenas de razonamiento largas o de profundidad innecesaria.

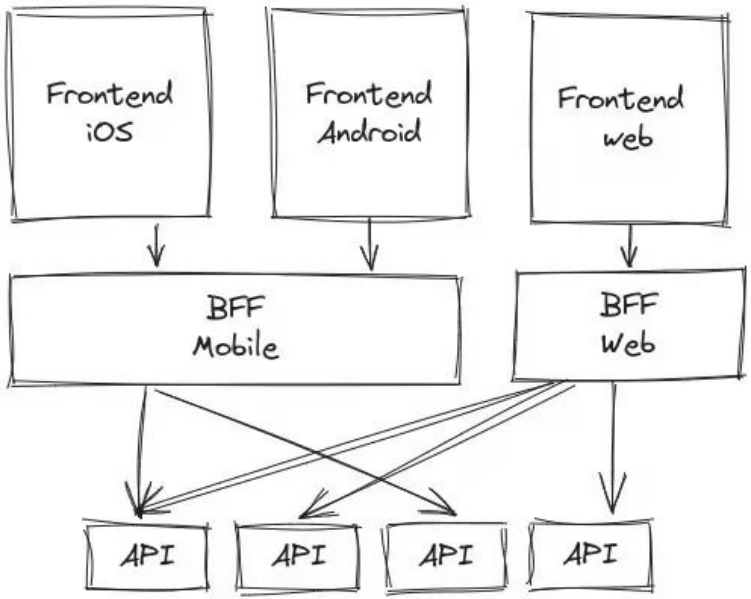
Tareas de ayuda rápida, tareas sencillas o repetitivas

Usar uno de estos modelos cuando queremos:

- Escribir o editar funciones pequeñas o código de utilidad.
- Formular preguntas rápidas sobre la sintaxis o el lenguaje.
- Crear prototipos de ideas con una configuración mínima.
- Obtener comentarios rápidos sobre mensajes simples o modificaciones.

Modelo	Por qué es una buena opción
o4-mini /o3-mini	Un modelo rápido y rentable para tareas de programación repetitivas o sencillas. Ofrece sugerencias claras y concisas.
Claude Sonnet 3.5	Equilibra respuestas rápidas con salida de calidad. Ideal para tareas pequeñas y explicaciones de código ligeras.
Gemini 2.0	Latencia extremadamente baja y soporte técnico multimodal (si está disponible). Ideal para comentarios rápidos e interactivos.

Tareas de razonamiento profundo y depuración



Estos modelos están diseñados para tareas que requieren **razonamientos paso a paso**, toma de decisiones complejas o reconocimiento de contexto elevado.

Funcionan bien cuando se necesita un análisis estructurados, generación de código completa o comprensión de varios archivos.

Tareas de razonamiento profundo y depuración

Usa uno de estos modelos si quieres:

- Depurar problemas complejos en varios archivos.
- Refactorizar bases de código grandes o interconectadas.
- Planear las características o la arquitectura entre capas.
- Sopesar las ventajas entre bibliotecas, patrones o flujos de trabajo.
- Analizar los registros, los datos de rendimiento o el comportamiento del sistema.

Modelo	Por qué es una buena opción
Claude Sonnet 3.7	Proporciona razonamiento híbrido que se adapta tanto a las tareas rápidas como al pensamiento más profundo.
Claude Sonnet 4	Mejora la versión 3.7 con finalizaciones más confiables y razonamiento más inteligente bajo presión.
Claude Opus 4.1	Modelo antrópico más completo. Mejora Claude Opus 4.
Claude Opus 4	Destaca en la estrategia, la depuración y la lógica multicapa.
Gemini 2.5 Pro	Razonamiento avanzado en contextos largos y análisis científicos o técnicos.

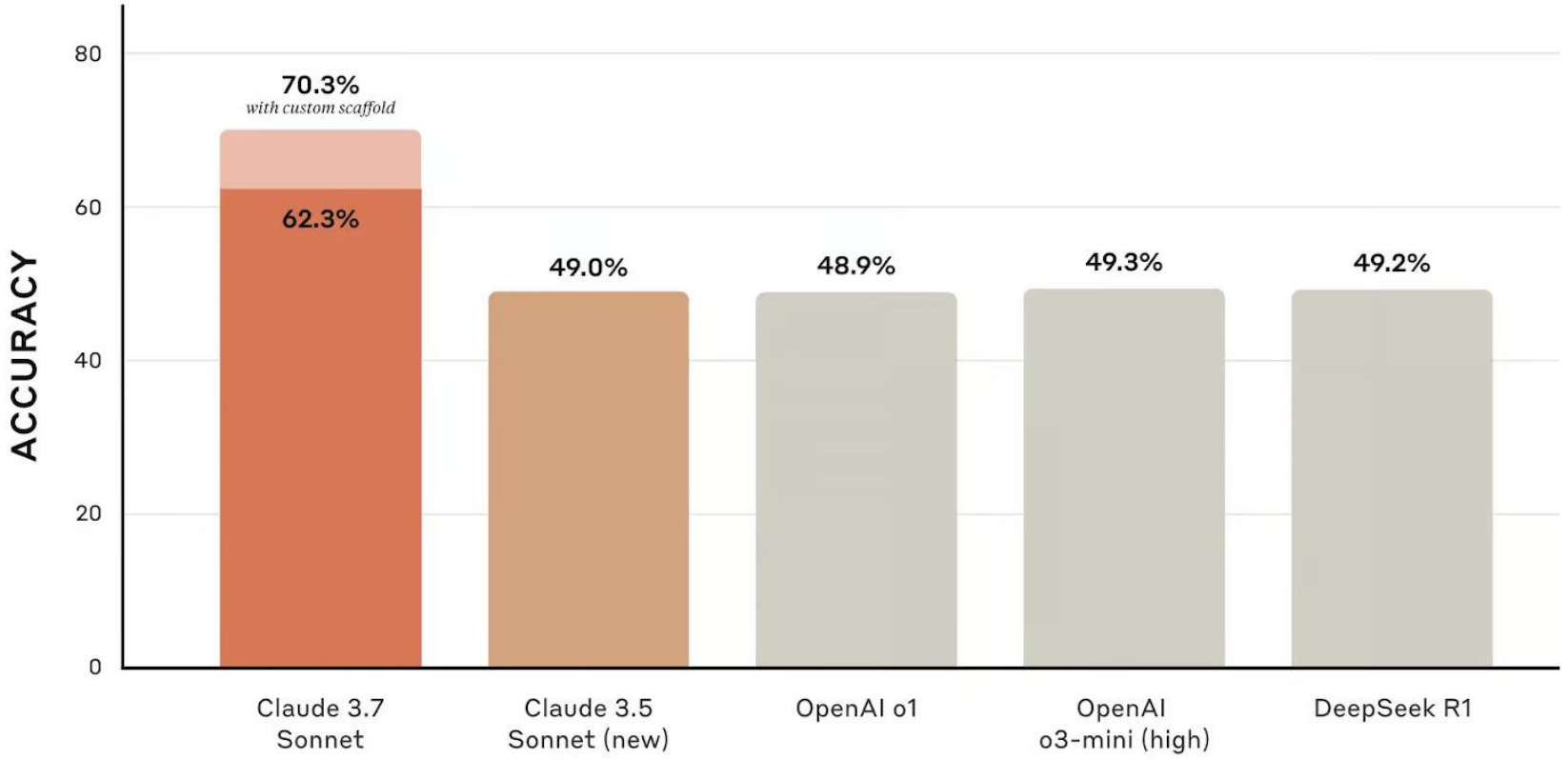
Claude 3.7 - Anthropic

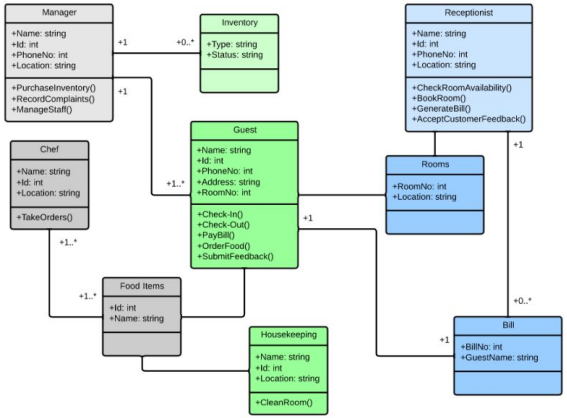
Destaca durante todo el ciclo de vida de desarrollo de software, desde el diseño inicial hasta las correcciones de errores, pasando por el mantenimiento y las optimizaciones.

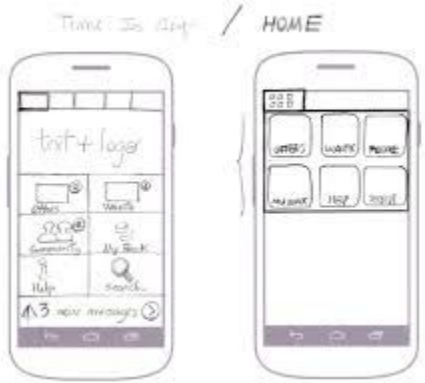
Es especialmente adecuado para la **refactorización** de varios archivos o la planificación arquitectónica, donde es importante comprender el **contexto** entre componentes.

Software engineering

SWE-bench verified







Trabajar con objetos visuales (diagramas, capturas de pantalla)

Cuando deseamos formular preguntas sobre capturas de pantalla, diagramas, componentes de la interfaz de usuario u otra entrada visual.

Estos modelos admiten entrada multimodal y son adecuados para el trabajo en el front-end o la depuración de objetos visuales.

Trabajar con objetos visuales (diagramas, capturas de pantalla)

Usar uno de estos modelos si deseas::

- Hacer preguntas o generar código sobre diagramas, capturas de pantalla o componentes de la interfaz de usuario.
- Obtener comentarios sobre borradores visuales o flujos de trabajo.
- Comprender el comportamiento de front-end desde el contexto visual.

Modelo	Por qué es una buena opción
GPT-4.1	Valor predeterminado confiable para la mayoría de las tareas de programación y escritura. Es rápido, preciso y compatible con la entrada multimodal para tareas de razonamiento visual. Funciona bien entre lenguajes y marcos.
Claude Opus 4	Modelo antrópico más completo. Destaca en la estrategia, la depuración y la lógica multicapa.
Claude Sonnet 4	Mejora la versión 3.7 con finalizaciones más confiables y razonamiento más inteligente bajo presión.
Gemini 2.0 Flash	Modelo rápido y multiplataforma optimizado para la interacción en tiempo real. Resulta útil para recibir comentarios sobre diagramas, prototipos visuales y diseños de interfaz de usuario.
Gemini 2.5 Pro	Razonamiento profundo y depuración, ideal para flujos de trabajo complejos de generación, depuración e investigación de código.

2 | ¿Cómo modificamos el modelo LLM?

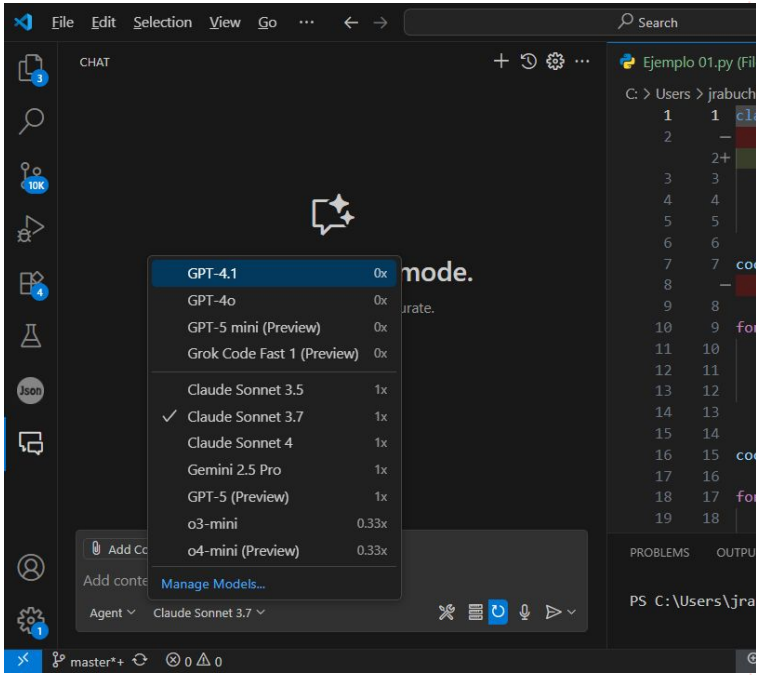
¿Cómo modificamos el modelo LLM en el Chat?

1 Abrir el Chat o Editor

Para abrir la vista de chat, haga clic en el icono de chat de la barra de actividad o presione Control+Comando+i (Mac)/Ctrl+Alt+i (Windows/Linux).

2 Seleccionar el modelo LLM

En la **parte inferior** de la vista de chat, seleccione el menú desplegable y, a continuación, haga clic en el modelo de IA que prefiera.



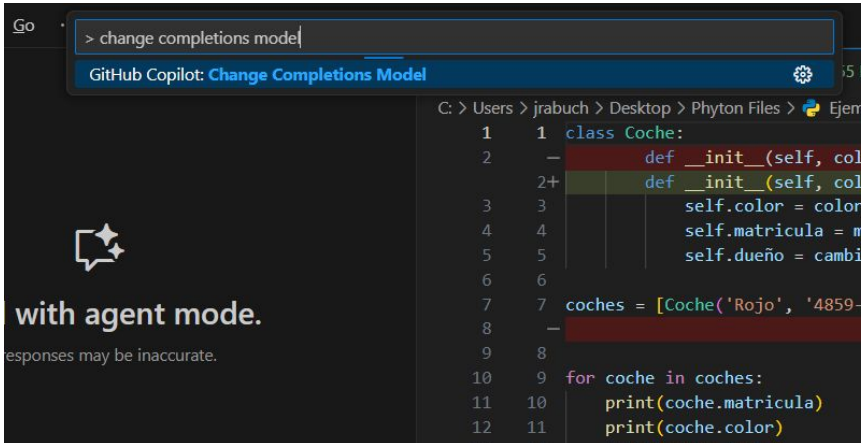
¿Cómo modificamos el modelo LLM para Sugerencias?

1 Abrir paleta de comandos

Para abrir la vista de chat, haga clic en el icono de chat de la barra de actividad o presione Comando+Shift+p (Mac)/Ctrl+Shift+p (Windows/Linux).

2 Seleccionar el modelo LLM

Escribir **change completions model** y selecciona el comando **"GitHub Copilot: Change Completions Model".**



Actividad en grupo

Armar un ejemplo y comparar su solución utilizando dos de los siguientes modelos LLM:

- GPT-4.1
- o3-mini
- Gemini 2.5 Pro
- Claude Sonnet

20:10 a 20:30 actividad

20:30 volvemos a la sala principal