

¡Ya casi empezamos!



Clase 12

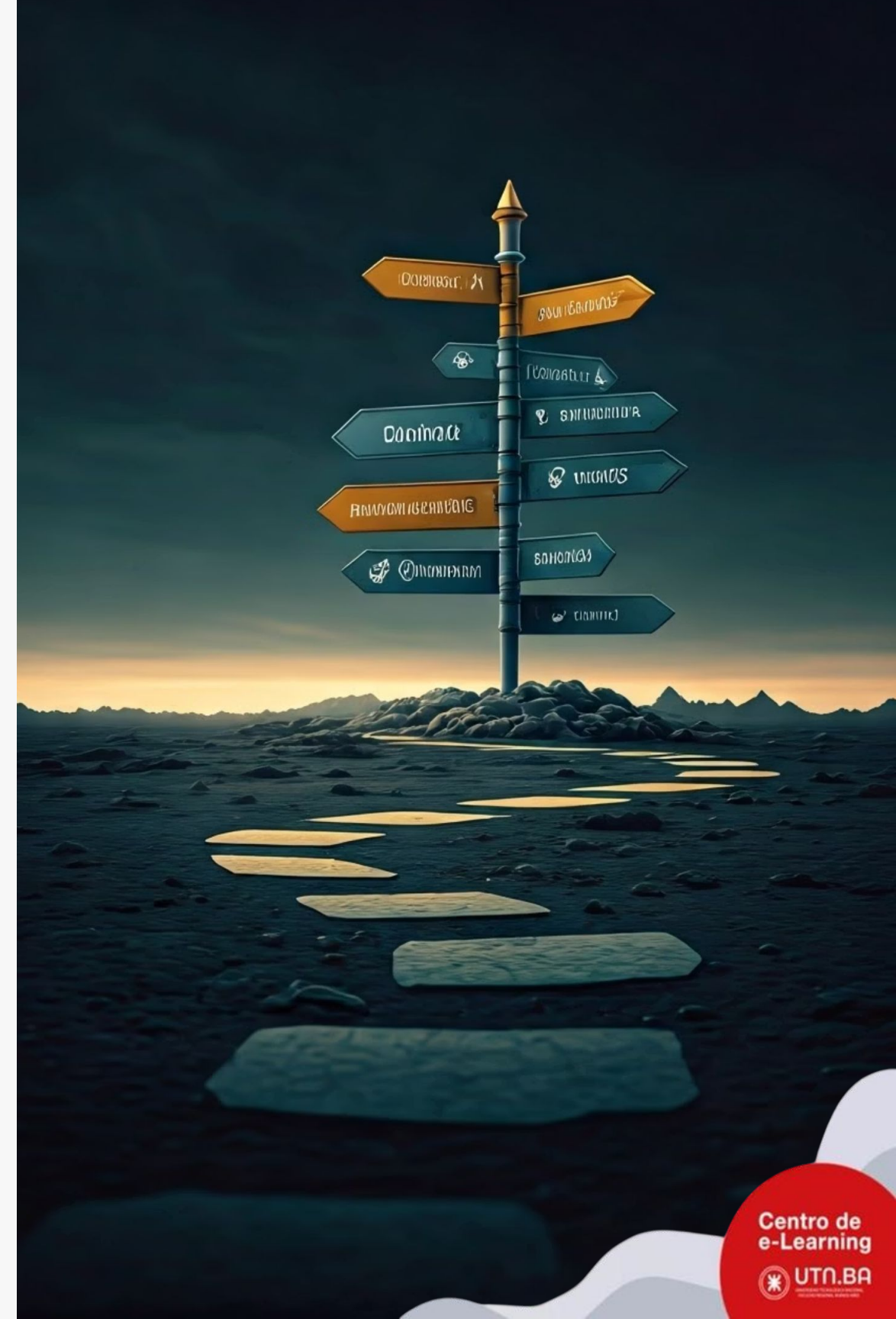
Integración de IA **en** **aplicaciones No Code**



Hoja de Ruta



1. Introducción y objetivos
2. ¿Dónde agrega valor la IA?
3. Consola de desarrolladores
4. Demo: consolas de los principales proveedores
5. ¿Qué es una API Key? - Creación de API Key
6. ¿Qué son los Tokens? ¿Cuánto pesan?
7. Ejercicio 1 - Pesar tus prompts
8. Ejercicio 2 - Tu primer presupuesto con IA
9. Cierre, preguntas y próxima clase



- ☒ Understand Key Concepts
- ☐ Apply Learned Skills
- ☐ Critical Thinking
- ☐ Collaborate

1. Objetivos



Entender dónde usar IA

Reconocer en qué pasos un modelo de IA puede potenciar una automatización



Conocer las plataformas

Familiarizarse con las consolas de desarrollador de OpenAI y Google



Identificar parámetros

Comprender los parámetros principales de configuración de LLMs



Cotizar soluciones con IA

Practicar con ejercicios para aprender a estimar los costos de un automatización con IA



2. Dónde la IA Agrega Valor Real

La inteligencia artificial no es solo una tecnología del futuro: ya está transformando operaciones cotidianas en organizaciones de todos los tamaños. Conocer dónde aplicarla marca la diferencia entre experimentos costosos y mejoras tangibles en productividad.

Clasificación Inteligente

Enrutamiento automático de tickets entre ventas y soporte técnico. Evaluación de CVs identificando candidatos aptos o no aptos, con justificación clara del criterio aplicado.

Resumen y generación

Conversión de emails extensos en 2-3 puntos accionables. Transformación de actas de reuniones en listas concisas de acuerdos alcanzados y tareas pendientes por completar.

Predicción Ligera

Asignación sugerida de prioridad (alta, media o baja) basada en señales simples del contexto, ayudando a tomar decisiones más rápidas y fundamentadas.

3. Consolas de desarrollador



Selector de modo/modelo

Permite elegir entre diferentes modelos del proveedor



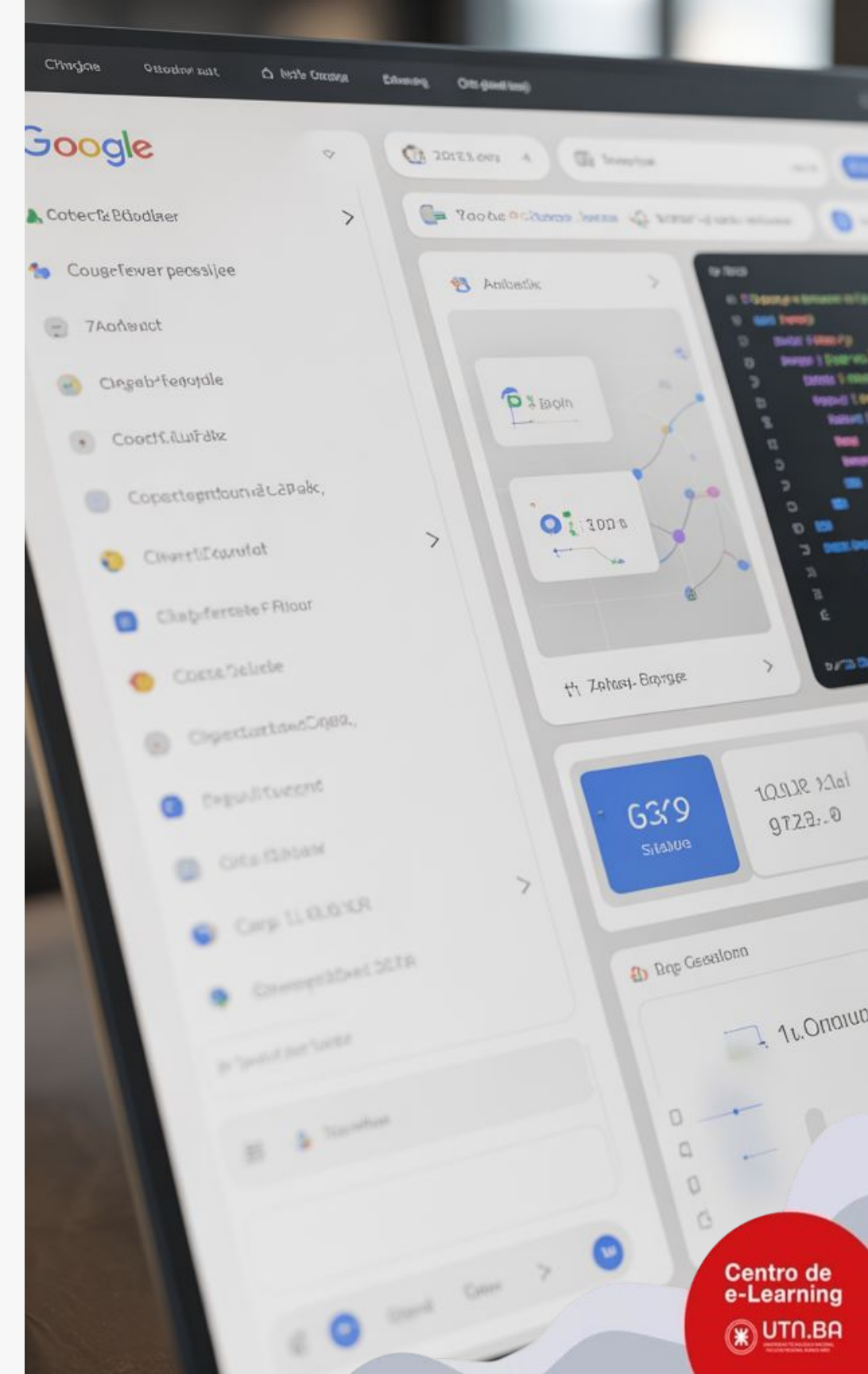
Caja de prompt

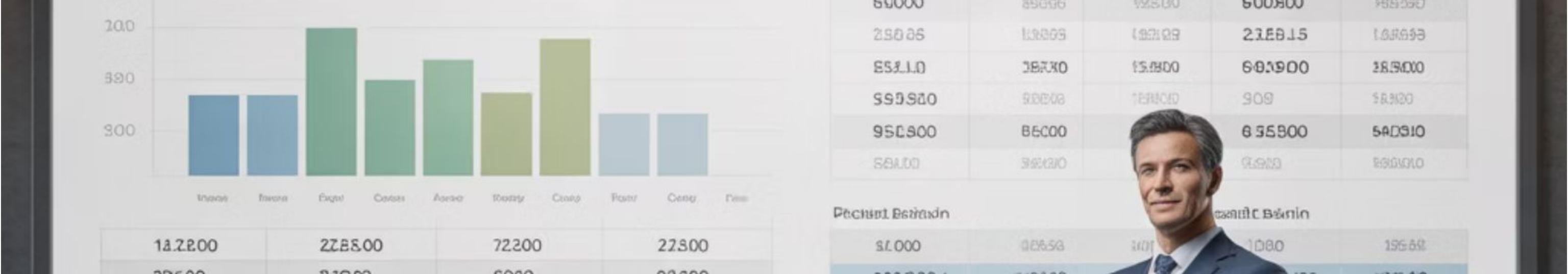
Área principal donde escribir las instrucciones para el modelo



Show code

Muestra el código necesario para implementar la misma llamada en tu aplicación





Parámetros de los modelos

Temperature

Controla la aleatoriedad de las respuestas. Valores más altos (como 0.8) hacen la salida más aleatoria, mientras que valores más bajos (como 0.2) la hacen más determinista.

Max tokens

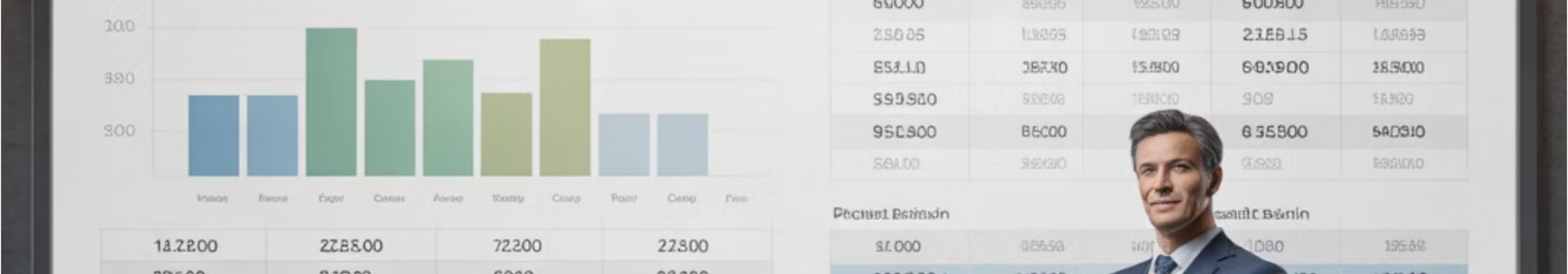
Limita la longitud de la respuesta generada. Un token es aproximadamente 4 caracteres o 3/4 de una palabra en inglés.

Top-p

Sampling con núcleo. El modelo considera los resultados de los tokens con probabilidad top_p. 0.1 significa que solo se consideran los tokens que comprenden el 10% superior de la masa de probabilidad.

Frequency/Presence penalty


Penaliza nuevas palabras basadas en su frecuencia existente en el texto (frequency) o simplemente si ya han aparecido (presence).





Parámetro	Ajuste “bajo”	Comportamiento	Ajuste “alto”	Comportamiento	Riesgo típico
Temperature	0.2	Predecible, estable	0.8	Creativo, variable	Divagar / inconsistencia
Top-p	0.3	Palabras comunes	1.0	Léxico más variado	Términos raros
Max tokens	40	Completo	8	Puede cortar	Respuestas truncadas
Frequency/Presence	0.0/0.0	Repite	0.9/0.9	Diversifica	Pérdida de tono si es excesivo

4. Demo: Consola de desarrolladores


OpenAI Console:


A blue gradient banner with the text "OpenAI Platform" in white.


 platform.openai.com
OpenAI Platform
Explore developer resources, tutorials, API docs, and dynamic examples to get the most out of OpenAI's platform.



Google AI Studio:

A dark blue banner with a star icon, the text "Google AI Studio", and a button that says "Build with AI models from Google DeepMind" and "Run".

 AI Studio
Google AI Studio
Google AI Studio is the fastest way to start building with Gemini, our next generation family of multimodal generative AI models.





5. ¿Qué es una API Key?



Identificador único

Cadena alfanumérica que identifica tu aplicación ante un servicio externo.



Autenticación y autorización

Verifica que la petición proviene de un cliente válido.



Facturación

Asocia el uso a tu cuenta para imputar el consumo.



Creación de API Keys en Google AI Studio

1

Acceso a la consola

Entra en console.cloud.google.com y selecciona tu proyecto.

2

Sección de credenciales

Ve a APIs & Services → Credentials.

3

Crear nueva clave

Haz clic en "Create credentials" → "API key".

4

Configurar restricciones

Establece límites por HTTP referrer, IP o aplicación.

6. ¿Qué son los Tokens?

Definición

Fragmentos de texto que el modelo utiliza para procesar entrada y generar salida.

Equivalencia

En inglés, 1 token $\approx \frac{3}{4}$ de palabra. Varía según idioma y vocabulario.

Importancia

El coste se basa en tokens de input + output. Conocerlos permite estimar presupuestos.

Tokens de entrada

Texto que envías al modelo:
instrucciones, contexto y datos.


Tokens de salida



Texto generado por el modelo.
Depende del "max_tokens".

¿Cuánto pesa un prompt?

“Tokenizadores”



Herramientas oficiales

OpenAI Platform

 platform.openai.com 

OpenAI Platform

Explore developer resources, tutorials, API docs, and dynamic examples to get the most out of OpenAI's platform.

 console.cloud.google.com 

Google Cloud console

7. Ejercicio 1:

"Pesar tus prompts"

1. Accede a <https://platform.openai.com/tokenizer>.
2. Obtén el texto de un **archivo input**, súbelo al tokenizer y anota el número de tokens que indica.
3. Pega tu **prompt** y anota el número de tokens que indica.
4. Ingresa a Chat GPT (o similar) ejecuta la prompt con el input.
5. Pega la **respuesta** obtenida en el tokenizer y registra también sus tokens.



Pricing de Modelos de IA

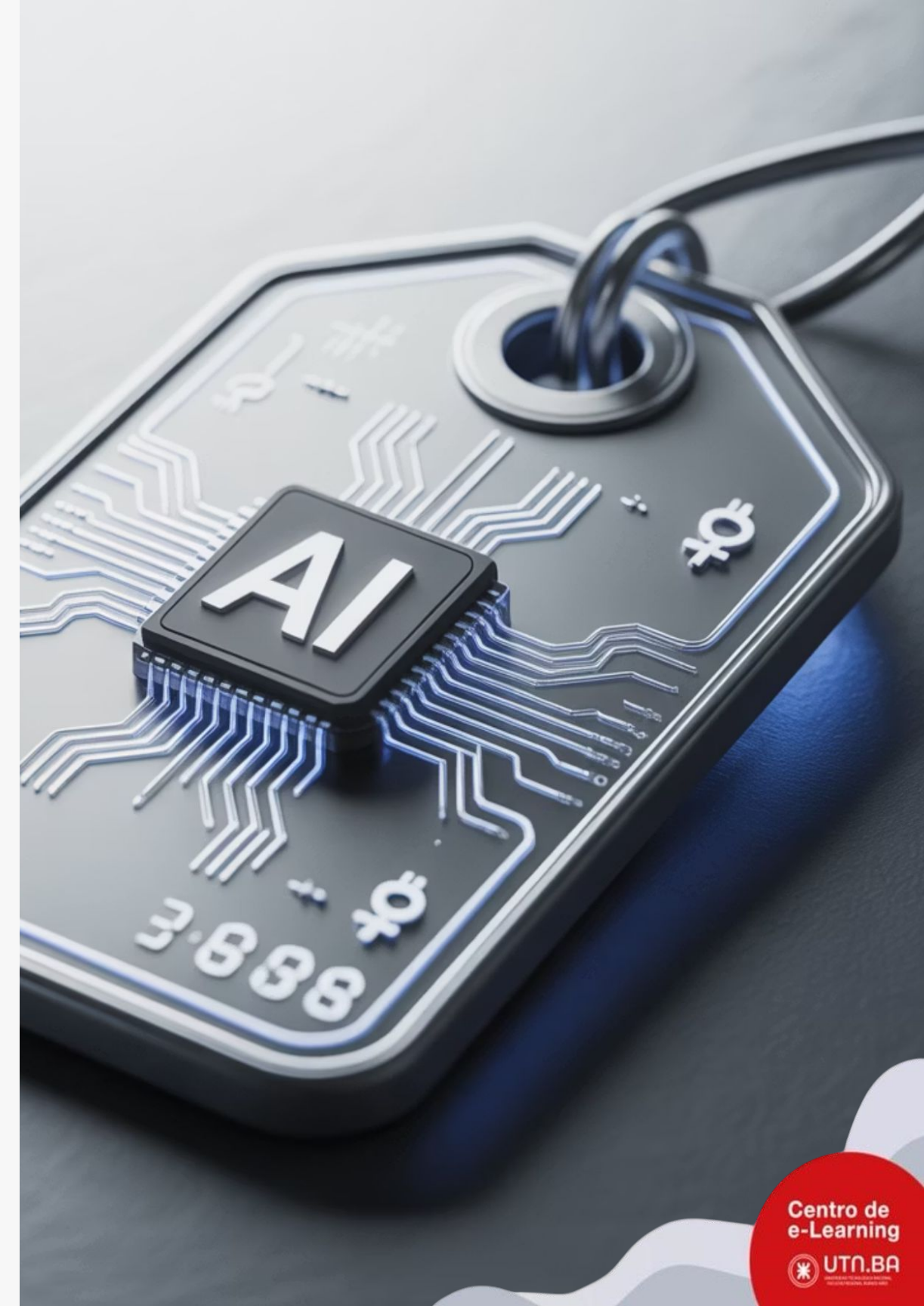
Los modelos de IA tienen estructuras de precios basadas principalmente en el uso de tokens. Comprender estas tarifas es fundamental para tomar decisiones informadas sobre qué modelo utilizar en cada caso de uso.

Factores de Influencia

Capacidades del modelo: Modelos más avanzados como GPT-5 cuestan más en su versión "thinking" (razonador).

Longitud del contexto: Modelos con ventanas de contexto más grandes (128k vs 8k tokens) tienen precios premium.

Volumen de uso: Algunos proveedores ofrecen descuentos por volumen mensual comprometido.



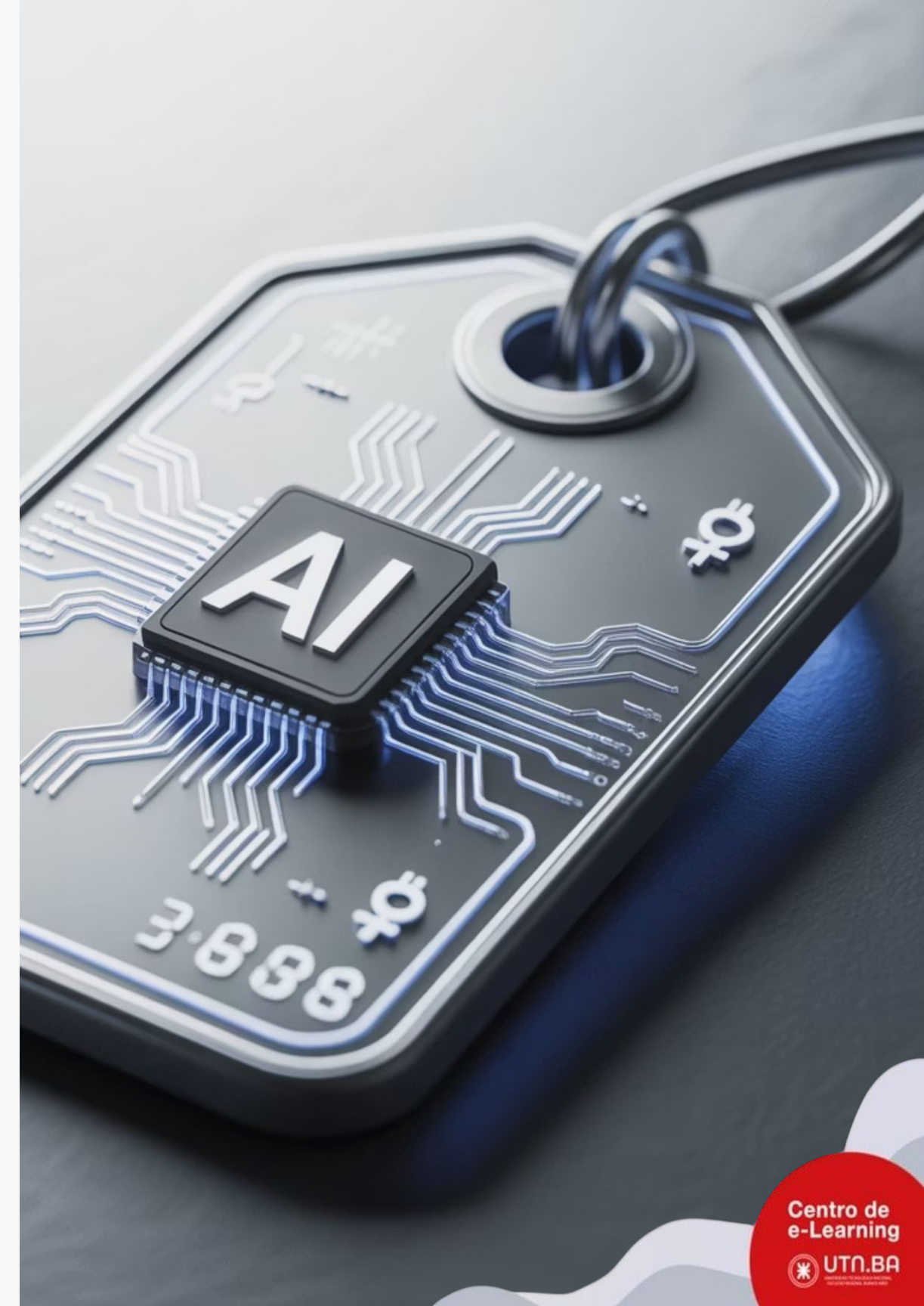
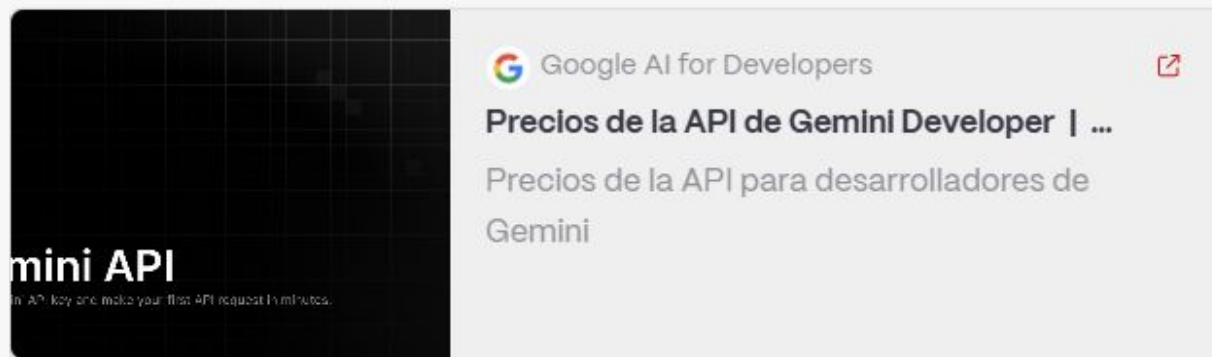
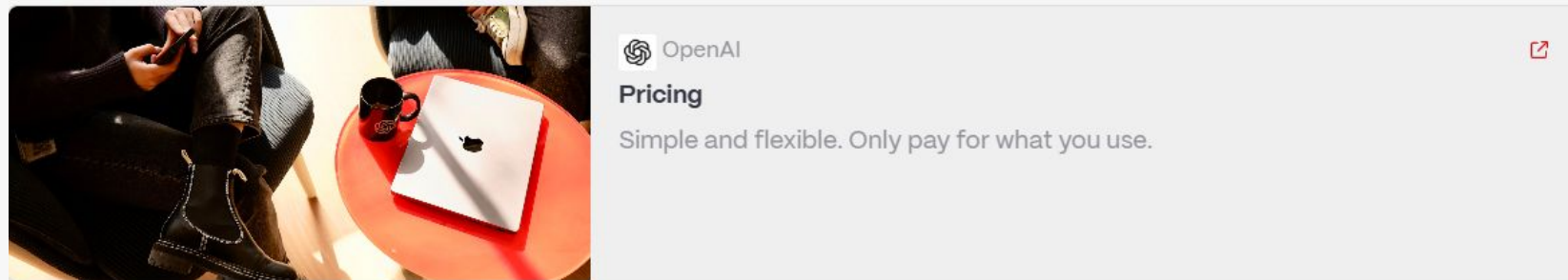
Pricing de Modelos de IA

Estructura de Precios

El pricing se presenta generalmente por cada 1,000,000 de tokens procesados, con costos diferenciados para tokens de entrada (prompt) y salida (respuesta generada).

Tokens de Entrada vs Salida

Los proveedores como OpenAI publican sus tarifas oficiales en su documentación. Típicamente, los tokens de salida cuestan más que los de entrada debido al mayor procesamiento requerido.



8. Presupuesto de Soluciones con IA

Calcular el costo real de una implementación de IA requiere considerar tanto tokens de entrada como de salida. Veamos un ejemplo práctico con números concretos.

Costo Input	Costo Output	Total por Llamada
$\text{Tokens} \times \text{Precio} = (500,000 / 1,000,000) \times \$2.00 = \text{\$1.00}$	$\text{Tokens} \times \text{Precio} = (200,000 / 1,000,000) \times \$4.00 = \text{\$0.80}$	$\text{Input} + \text{Output} = \$1.00 + \$0.80 = \text{\$1.80}$

Proyección Mensual

Con 10 llamadas al mes, el costo total sería: $\$1.80 \times 10 = \text{\$18.00 mensuales}$

Este método de cálculo te permite escalar fácilmente para cualquier volumen de uso. Multiplica el costo por llamada por tu estimación realista de ejecuciones mensuales para obtener tu presupuesto base.

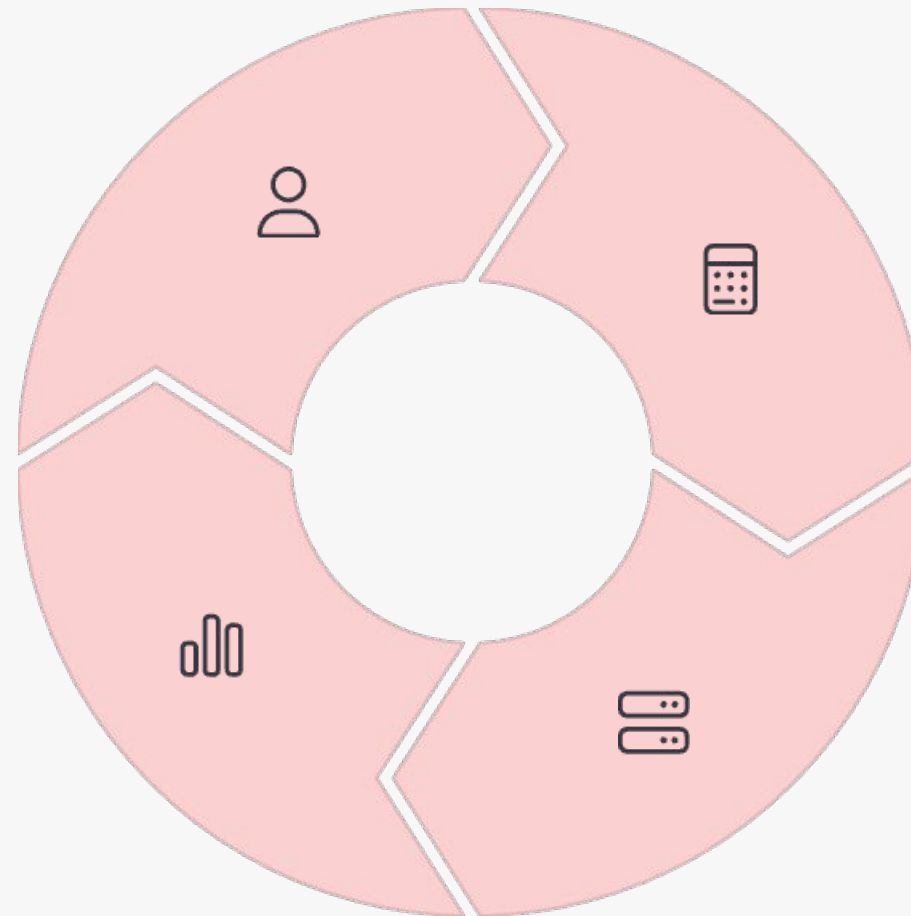
Cálculo de Presupuesto

Estimar volumen

Número de usuarios y peticiones
por día

Escenarios

Conservador, medio y optimista



Aplicar fórmula

Calcular coste diario y mensual

Costes adicionales

Networking, storage y licencias

9. Ejercicio 2

Cotizando soluciones de IA

1. Tomar los resultados del Ejercicio 1 (tokens totales por ejecución)
2. Seleccionar un modelo acorde:
<https://platform.openai.com/docs/models> ó
<https://ai.google.dev/gemini-api/docs/models>
3. Consultar el pricing del modelo: <https://openai.com/api/pricing/> ó
<https://ai.google.dev/gemini-api/docs/pricing>
4. Calcular el **coste por ejecución**: $(\text{Tokens totales} / 1\,000) \times \text{Precio}$
5. Estimar un **coste mensual** para 1 000 ejecuciones.



¡Nos vemos en la próxima clase!

Seguiremos con...

Creación de asistentes virtuales

¡Gracias!