Un enfoque innovador para interactuar eficazmente con la inteligencia artificial

Generador de Prompts: Tutorial de Uso

Eddie Rivera

Tabla de contenido

01

Tutorial de Uso: Generador de Prompts Estructurados

Tutorial de Uso

Esta aplicación te ayuda a crear prompts claros y efectivos para obtener mejores resultados de IA. Su diseño está orientado a búsqueda general y a la administración de empresas.

En un mundo donde las tecnologías basadas en modelos de lenguaje (LLM, por sus siglas en inglés) están transformando la interacción entre personas e inteligencia artificial, la calidad de la comunicación mediante prompts se ha convertido en un factor crucial. El lenguaje, como medio principal para entendernos, no solo define la claridad y precisión de las respuestas generadas por las IA, sino que también determina la efectividad y el impacto de su uso en diversos contextos.

La redacción de prompts deficientes puede generar respuestas inexactas, lo que incrementa la necesidad de iteraciones adicionales para alcanzar los resultados deseados. Esto no solo afecta la productividad, sino que también repercute directamente en los recursos utilizados. Según estudios recientes, empresas que dependen de IA para tareas como atención al cliente o análisis de datos pueden enfrentar hasta un 25% de aumento en costos operativos debido a solicitudes mal formuladas. Por ejemplo, en una organización promedio que emplea IA con un gasto anual de $500,000 en infraestructura y tiempo de procesamiento, mejorar la precisión en los prompts podría ahorrar hasta $125,000 al año en recursos y evitar miles de horas laborales desperdiciadas.

Por ello, este tutorial busca no solo sensibilizar sobre la importancia de estructurar correctamente los prompts, sino también proporcionar herramientas prácticas para optimizar la interacción con LLM. Una comunicación efectiva no solo reduce costos, sino que maximiza el potencial de las IA, impulsando resultados que beneficien tanto a individuos como a organizaciones.

Si lo que busca son prompts para generar imágenes, esta no es la aplicación adecuada. Próximamente se estará presentando una versión para estos fines.
A continuación, se explica cada sección y campo del formulario utilizado en la app, con ejemplos y consejos.

02

Rol de la IA

Una estructura clara en los prompts es esencial para maximizar la utilidad de la interacción con la IA. Esto no solo mejora la calidad de las respuestas, sino que también garantiza que estas sean coherentes y alineadas con las expectativas del usuario. Al combinar una tarea principal bien definida, una solicitud específica y un contexto detallado, los usuarios pueden explorar las capacidades de la IA de manera más efectiva.

Ejemplo:

“Analizar las implicaciones éticas de la IA en educación universitaria para poder diseñar planes curriculares efectivos.”

¿Qué es?

El “Rol de la IA” define la personalidad, función o punto de vista desde el cual la inteligencia artificial responderá a tu solicitud. Esto ayuda a contextualizar la respuesta y a obtener resultados más alineados con tus expectativas. Por ejemplo, puedes pedirle a la IA que actúe como un experto en un área específica, como un profesor, un consultor, un programador, o incluso como un personaje histórico o literario. Elegir el rol correcto puede influir en el tono, el nivel de detalle y el tipo de información que la IA proporcionará.

Opciones típicas:

Sistema: La IA actúa como el sistema que controla la conversación, funciona como el núcleo del que organiza y controla la interacción. En este rol, la IA procesa las solicitudes, administra la información, y regula la estructura de las conversaciones para garantizar un flujo lógico. Es ideal para configuraciones donde se requiere una experiencia controlada o automatizada.

Ejemplo de redacción:

- “Eres un sistema que regula las conversaciones en un entorno de aprendizaje virtual. Tu tarea es organizar las interacciones entre estudiantes y profesores, garantizar que las preguntas sean claras y que las respuestas sean relevantes y bien estructuradas, mientras aseguras el cumplimiento de las normas del aula virtual.”
- “Eres un asesor financiero especializado en inversiones sostenibles.”
- “Actúa como un profesor universitario de literatura latinoamericana.”

Usuario: La IA responde como si fuera el usuario, asume el papel de un interlocutor. Este enfoque es útil para simular escenarios en los que se desea analizar la perspectiva del usuario o explorar cómo un usuario podría interactuar en distintas situaciones. Por ejemplo, en pruebas de usabilidad, la IA puede actuar como un usuario típico para evaluar un sistema.

Ejemplo de redacción:

“Eres un usuario que interactúa con un sistema de aprendizaje virtual. Tu tarea es simular cómo un estudiante típico realizaría preguntas sobre un curso de programación, solicitando aclaraciones sobre temas como bucles, funciones y estructuras de datos.”

Asistente: La IA asume el papel de un colaborador experto en un campo específico, ofreciendo orientación, análisis o soluciones detalladas y personalizadas. Este rol es ideal cuando se busca aprovechar el conocimiento acumulado de la IA para resolver problemas complejos, proporcionar asesoramiento especializado o asistir en tareas técnicas y creativas.

Ejemplo de redacción:

“Eres un asistente experto en análisis de datos. Tu tarea es ayudar a un equipo de marketing a interpretar los resultados de una campaña publicitaria, proporcionando insights clave sobre métricas como el retorno de inversión (ROI), la tasa de conversión y el engagement en redes sociales. Además, debes sugerir estrategias basadas en los datos para optimizar futuras campañas.”

Asistente: La IA actúa como un ayudante o experto en un área específica. Esto incluye responder preguntas, brindar asesoramiento o realizar tareas de apoyo en base a conocimientos especializados. Es ideal para escenarios donde se necesita una guía práctica, análisis detallados o generación de contenido basado en habilidades avanzadas.

Ejemplo de redacción:

“Eres un asistente experto en análisis de datos. Tu tarea es ayudar a un equipo de marketing a interpretar los resultados de una campaña publicitaria, proporcionando insights clave sobre métricas como el retorno de inversión (ROI), la tasa de conversión y el engagement en redes sociales. Además, debes sugerir estrategias basadas en los datos para optimizar futuras campañas.”

03

2. Estructura del Prompt

Cuando se redacta un prompt para una IA, es fundamental que su estructura sea precisa y clara para maximizar la eficacia y relevancia de las respuestas obtenidas. Una buena estructura de prompt no solo facilita la interacción con el sistema, sino que también permite a los usuarios obtener resultados que cumplan con sus expectativas. A continuación, se describen los elementos clave de una estructura de prompt bien elaborada:

a. Tarea Principal

¿Qué es?

La “Tarea Principal” es la acción central que deseas que la IA realice. Es el objetivo principal del prompt y debe ser lo más claro y específico posible.

Una buena tarea principal ayuda a enfocar la respuesta y evita ambigüedades.

Este componente define el propósito general de la interacción con la IA. Es una descripción concisa de la acción o análisis que se espera de la IA, y debe reflejar el objetivo principal del usuario. La tarea principal establece el marco de referencia para todo el prompt y orienta el enfoque de la respuesta hacia un área específica. Puede ser formulada en términos amplios o específicos, dependiendo del nivel de detalle que se requiera. Tiene que comenzar con un verbo o comando específico.

Ejemplo:

- "Identificar tendencias emergentes en el mercado tecnológico."
- "Analizar las implicaciones éticas de la IA en educación universitaria."
- "Redactar una carta de presentación para una solicitud de empleo."

Técnicas de prompting para definir la tarea principal

1. General Prompting / Zero-shot

A. Técnica: General Prompting / Zero-shot.

B. Explicación: Es el tipo de prompt más sencillo. Solo proporciona una descripción de una tarea y algo de texto para que el modelo de lenguaje grande (LLM) comience. El nombre "zero-shot" significa que no se proporcionan ejemplos en el prompt. La entrada puede ser una pregunta, el inicio de una historia o instrucciones.

C. Uso principal: Iniciar una tarea, como clasificar reseñas de películas.

D. Ejemplo:

Prompt: "Clasifica reseñas de películas como POSITIVA, NEUTRAL o NEGATIVA. Reseña: 'Her' es un estudio inquietante que revela la dirección a la que se dirige la humanidad si se permite que la IA siga evolucionando, sin control. Desearía que hubiera más películas como esta obra maestra. Sentimiento:"

Output: POSITIVE

E. Limitación: Si la técnica zero-shot no produce los resultados deseados, se sugiere proporcionar demostraciones o ejemplos en el prompt, lo que lleva a las técnicas "one-shot" y "few-shot". Para tareas que no requieren creatividad, como la clasificación, la temperatura del modelo debe ser baja.

2. One-shot & Few-shot Prompting

A. Técnica: One-shot & Few-shot Prompting.

B. Explicación: Al crear prompts para modelos de IA, es útil proporcionar ejemplos para ayudar al modelo a entender lo que se le pide.

Un prompt one-shot proporciona un solo ejemplo. La idea es que el modelo tenga un ejemplo que pueda imitar para completar mejor la tarea.

Un prompt few-shot proporciona múltiples ejemplos al modelo. Este enfoque muestra al modelo un patrón que debe seguir, lo que aumenta la probabilidad de que el modelo siga ese patrón en su generación.

C. Uso principal: Son especialmente útiles cuando se desea guiar al modelo hacia una estructura de salida o patrón específico. Para tareas de clasificación con few-shot, se recomienda mezclar las clases de respuesta en los ejemplos para evitar el sobreajuste.

D. Ejemplo: Parsear pedidos de pizza a formato JSON. Los ejemplos muestran cómo el modelo debe estructurar la salida JSON para diferentes tipos de pedidos, incluyendo casos con ingredientes divididos por mitades.

E. Limitación: La cantidad de ejemplos necesarios depende de la complejidad de la tarea, la calidad de los ejemplos y las capacidades del modelo. Se recomienda al menos de tres a cinco ejemplos para few-shot, pero la limitación de longitud de entrada del modelo puede requerir menos. Un pequeño error en los ejemplos puede confundir al modelo y resultar en una salida no deseada.

3. System, Contextual y Role Prompting

Estas tres técnicas se utilizan para guiar cómo los LLM generan texto, enfocándose en diferentes aspectos, aunque pueden superponerse.

3.1. System Prompting

A. Técnica: System Prompting.

B. Explicación: Establece el contexto general y el propósito principal para el modelo de lenguaje, definiendo la "gran imagen" de lo que el modelo debería estar haciendo, como traducir un idioma o clasificar una reseña.

C. Uso principal: Generar salidas que cumplan con requisitos específicos, como un formato de salida particular (ej. JSON). También es muy útil para la seguridad y la toxicidad, permitiendo añadir instrucciones para controlar la salida del modelo.

D. Ejemplo:

Clasificar reseñas de películas y devolver solo la etiqueta en mayúsculas.

Clasificar reseñas de películas y devolver la salida en formato JSON válido según un esquema específico.

E. Limitación: La generación de objetos JSON requiere significativamente más tokens que el texto plano, lo que puede aumentar el tiempo de procesamiento y los costos.

3.2. Role Prompting

A. Técnica: Role Prompting.

B. Explicación: Implica asignar un rol o identidad específica al modelo de IA generativa (ej. editor de libros, guía de viajes, orador motivacional). Esto ayuda al modelo a generar respuestas que son consistentes con el rol asignado y su conocimiento y comportamiento asociados.

C. Uso principal: Definir una perspectiva de rol para un modelo de IA le proporciona un modelo del tono, estilo y experiencia específica que se busca para mejorar la calidad, relevancia y efectividad de la salida.

D. Ejemplo:

"Actúa como guía de viajes para proporcionar sugerencias de lugares para visitar."

Modificar el prompt para incluir un estilo humorístico e inspirador al actuar como guía de viajes.

3.3. Contextual Prompting

A. Técnica: Contextual Prompting.

B. Explicación: Proporciona detalles específicos o información de fondo relevante para la conversación o tarea actual. Ayuda al modelo a comprender los matices de lo que se le pide y a adaptar la respuesta en consecuencia. Es altamente específico para la tarea o entrada actual y es dinámico.

C. Uso principal: Asegurar que las interacciones con la IA sean fluidas y eficientes, permitiendo que el modelo entienda más rápidamente la solicitud y genere respuestas más precisas y relevantes.

D. Ejemplo: Sugerir 3 temas de artículos para un blog de juegos de arcade retro de los 80, proporcionando el contexto de que el modelo está escribiendo para dicho blog.

4. Step-back Prompting

A. Técnica: Step-back Prompting.

B. Explicación: Mejora el rendimiento al pedirle al LLM que primero considere una pregunta general relacionada con la tarea específica, y luego usar la respuesta a esa pregunta general en un prompt subsiguiente para la tarea específica. Este "paso atrás" permite al LLM activar conocimiento de fondo relevante y procesos de razonamiento antes de intentar resolver el problema específico.

C. Uso principal: Generar respuestas más precisas e perspicaces al considerar principios más amplios y subyacentes. Fomenta el pensamiento crítico y la aplicación de conocimiento de nuevas formas. Puede ayudar a mitigar sesgos al enfocarse en principios generales en lugar de detalles específicos.

D. Ejemplo: En lugar de pedir directamente una historia para un videojuego, primero se le pregunta al modelo cuáles son 5 configuraciones clave que contribuyen a una historia atractiva para un juego de disparos en primera persona. Luego, una de estas configuraciones se utiliza como contexto para pedir la historia final, resultando en una trama más interesante y específica.

E. Limitación: Si la temperatura se establece en un valor alto (ej. 1) sin aplicar la técnica de "step-back", la salida puede ser "aleatoria y genérica".

5. Chain of Thought (CoT)

Tutorial de Uso

A. Técnica: Chain of Thought (CoT) Prompting.

B. Explicación: Mejora las capacidades de razonamiento de los LLM al generar pasos de razonamiento intermedios. Esto ayuda al LLM a producir respuestas más precisas. Puede combinarse con few-shot prompting para obtener mejores resultados en tareas complejas que requieren razonamiento.

C. Uso principal: Resolver tareas que requieren razonamiento, como problemas matemáticos. Es útil para la generación de código (desglosando solicitudes en pasos), la creación de datos sintéticos, y cualquier tarea que pueda resolverse "dialogando" los pasos. Ofrece interpretabilidad al permitir ver los pasos de razonamiento del LLM, lo que facilita la identificación de fallos. Mejora la robustez entre diferentes versiones de LLM.

D. Ejemplo:

Prompt inicial (sin CoT): "Cuando yo tenía 3 años, mi pareja tenía 3 veces mi edad. Ahora, tengo 20 años. ¿Cuántos años tiene mi pareja?"

Output inicial (incorrecto): 63 years old

Prompt con CoT: Se añade "Let's think step by step."

Output con CoT: El modelo detalla cada paso de razonamiento hasta llegar a la respuesta correcta de 26 years old. Se muestra también un ejemplo de CoT con single-shot para guiar el modelo con un patrón de razonamiento específico.

6. Self-consistency

A. Técnica: Self-consistency.

B. Explicación: Combina el muestreo y la votación mayoritaria para generar diversas rutas de razonamiento y seleccionar la respuesta más consistente. Se proporciona el mismo prompt al LLM varias veces, utilizando una configuración de temperatura alta para fomentar la generación de diferentes perspectivas de razonamiento. Luego, se extrae la respuesta de cada generación y se elige la que aparece con más frecuencia.

C. Uso principal: Mejorar la precisión y coherencia de las respuestas generadas por los LLM, especialmente en tareas donde el razonamiento puede ser complejo o susceptible a ambigüedades (ej. debido al tono, elección de palabras o sarcasmo en la entrada).

D. Ejemplo: Un sistema de clasificación de correo electrónico que clasifica un email como "IMPORTANTE" o "NO IMPORTANTE". El email contiene sarcasmo que puede confundir al LLM. Se muestra cómo el modelo puede generar diferentes explicaciones y conclusiones ("IMPORTANT" vs. "NOT IMPORTANT") en múltiples intentos. La respuesta final más consistente se obtiene tomando la más común ("IMPORTANT").

7. Tree of Thoughts (ToT)

A. Técnica: Tree of Thoughts (ToT).

B. Explicación: Generaliza el concepto de Chain of Thought (CoT) al permitir que los LLM exploren múltiples rutas de razonamiento diferentes simultáneamente, en lugar de seguir una única cadena lineal de pensamiento. Mantiene un "árbol de pensamientos", donde cada pensamiento es una secuencia coherente de lenguaje que sirve como paso intermedio hacia la solución de un problema. El modelo puede ramificarse desde diferentes nodos del árbol para explorar distintas rutas de razonamiento.

C. Uso principal: Particularmente adecuada para tareas complejas que requieren exploración.

D. Ejemplo: La fuente proporciona una visualización conceptual que ilustra la diferencia entre CoT (lineal) y ToT (ramificada) [209, Figura 1]. No se presenta un prompt de ejemplo ejecutable.

8. ReAct (reason & act)

A. Técnica: ReAct (reason & act).

B. Explicación: Es un paradigma que permite a los LLM resolver tareas complejas mediante el razonamiento en lenguaje natural combinado con el uso de herramientas externas (como búsqueda en la web o intérpretes de código). Imita cómo los humanos operan, razonando verbalmente y tomando acciones para obtener información. Funciona a través de un bucle de pensamiento-acción: el LLM razona, genera un plan de acción, realiza las acciones, observa los resultados y actualiza su razonamiento y plan hasta alcanzar una solución. Se considera un primer paso hacia el modelado de agentes.

C. Uso principal: Resolver tareas complejas que requieren interactuar con el mundo exterior o recuperar información que no está en los parámetros internos del modelo. Se desempeña bien en diversos dominios.

D. Ejemplo: Determinar cuántos hijos tienen los miembros de la banda Metallica. El modelo realiza una serie de búsquedas externas (simuladas con SerpAPI) para cada miembro, observa los resultados de las búsquedas y encadena sus pensamientos para llegar a la respuesta final (10 hijos).

9. Automatic Prompt Engineering (APE)

A. Técnica: Automatic Prompt Engineering (APE).

B. Explicación: Un método que automatiza la creación de prompts. Implica que un modelo genere prompts, los evalúe (usando métricas como BLEU o ROUGE), posiblemente altere los buenos y repita el proceso. Esto no solo reduce la necesidad de intervención humana, sino que también mejora el rendimiento del modelo en diversas tareas.

C. Uso principal: Entrenar chatbots para entender diversas formas en que los clientes podrían expresar una solicitud. También para generar variantes de instrucciones o prompts para aplicaciones de software.

D. Ejemplo: Generar 10 variantes de la frase de un pedido de camiseta: "Una camiseta de Metallica talla S", manteniendo el mismo significado, para entrenar un chatbot de una tienda web de mercancía de bandas.

Técnicas de prompting y modelos LLM

Esta tabla presenta ejemplo de alineamiento entre el tipo de prompt que se desea utilizar y el modelo LLM más práctico para su uso.

Técnica de Prompting

ChatGPT (OpenAI)

Gemini (Google)

Claude (Anthropic)

1. General Prompting / Zero-shot

GPT-3.5, GPT-4, GPT-4.1 mini

Gemini 1.0, Gemini 1.5 Flash

Claude 3 Haiku, Claude 3 Sonnet

2. One-shot & Few-shot Prompting

GPT-3.5, GPT-4, GPT-4.5

Gemini 1.5 Pro, Gemini 1.0 Pro

Claude 3 Sonnet, Claude 3 Opus

3.1. System Prompting

GPT-4, GPT-4.5

Gemini 1.5 Pro

Claude 3 Opus

3.2. Role Prompting

GPT-4, GPT-4.5

Gemini 1.5 Pro

Claude 3 Sonnet, Claude 3 Opus

3.3. Contextual Prompting

GPT-4, GPT-4.5

Gemini 1.5 Pro, Gemini 1.5 Flash

Claude 3 Sonnet, Claude 3 Opus

4. Step-back Prompting

GPT-4.5

Gemini 1.5 Pro

Claude 3 Opus

5. Chain of Thought (CoT)

GPT-4, GPT-4.5

Gemini 1.5 Pro

Claude 3 Sonnet, Claude 3 Opus

6. Self-consistency

GPT-4.5

Gemini 1.5 Pro

Claude 3 Opus

7. Tree of Thoughts (ToT)

GPT-4.5 (con agentes / herramientas externas)

Gemini 1.5 Pro (con agentes o combinaciones)

Claude 3 Opus (con herramientas externas y planificación)

b. Pregunta o Solicitud Principal

¿Qué es?

Aquí se formula la pregunta o solicitud específica que el usuario desea que la IA aborde. Este apartado debe ser claro y directamente relacionado con la tarea principal, orientando a la IA hacia un aspecto particular que se desea explorar. Es crucial que la pregunta sea concreta para evitar respuestas ambiguas o demasiado generales.

Ejemplo:

- "¿Cuáles son las tecnologías más prometedoras en inteligencia artificial para los próximos cinco años?"
- "¿Cómo pueden las universidades implementar políticas éticas de IA?"

c. Descripción Detallada

¿Qué es?
Información adicional para contextualizar la tarea.

Este elemento proporciona contexto adicional que enriquece la comprensión de la IA sobre la tarea y asegura que la respuesta sea relevante y completa. Incluye detalles que pueden abarcar restricciones, consideraciones, enfoques o categorías específicas que deben ser tomadas en cuenta. La descripción detallada permite que la IA adapte su respuesta para cumplir con el nivel de especificidad requerido.

Ejemplo:

- “El análisis debe incluir referencias a publicaciones académicas y datos de mercado recientes, con un enfoque en aplicaciones comerciales.”
- “El análisis debe considerar aspectos pedagógicos, técnicos y de privacidad.”

d. Contexto General

¿Qué es?
El contexto general es la situación, entorno o antecedentes en los que se enmarca tu solicitud. Proporcionar contexto permite a la IA adaptar la respuesta a circunstancias particulares, como el país, la industria, el nivel educativo, etc.

Ejemplo:

"En el contexto de una universidad latinoamericana con recursos limitados y alta diversidad cultural."

e. Aportes del Usuario

¿Qué es?
Son datos, información o recursos adicionales que tú, como usuario, aportas para enriquecer la respuesta de la IA.
Esto puede incluir estadísticas, resultados de encuestas, documentos, enlaces, etc.

Ejemplo:

"Proporciono datos de encuestas recientes sobre adopción de IA en universidades."

f. Consideraciones del Contexto

¿Qué es?
Son restricciones, condiciones especiales o factores que la IA debe tener en cuenta al generar la respuesta.
Esto puede incluir limitaciones legales, culturales, técnicas o de cualquier otro tipo.

Ejemplo:

"Considerar regulaciones locales y diferencias culturales en la adopción tecnológica."

g. Expectativa del Resultado

¿Qué es?
Describe el formato, nivel de profundidad o tipo de resultado que esperas recibir.
Esto ayuda a la IA a ajustar la extensión, el detalle y la estructura de la respuesta.

Ejemplo:

Uso esperado del promt.

Creativo

Crea contenido único como texto, imágenes, audio y videos, a menudo estilizados para cumplir instrucciones específicas, p. ej. “Escribe al estilo de Charles Dickens.”

Instruccional

Ofrece orientación, como servicio al cliente. Un prompt instructivo podría ser: “Explica los pasos que un cliente necesita seguir para actualizar su cuenta.”

Informativo

Reúne y sintetiza información, ya sea de datos de entrenamiento o recursos subidos como archivos PDF. Usa prompts informativos de IA para obtener resúmenes o perspectivas.

Lista

Compila listas, desde simples hasta más detalladas y estructuradas. Por ejemplo, pide a la IA que genere una lista de títulos de blogs o ideas de temas.

Interactivo

Imita conversaciones reales, como entrevistas simuladas o escenarios de juego de roles.

Razonamiento

Analiza información y saca conclusiones. Por ejemplo, utiliza un prompt para identificar un mercado objetivo a partir de datos de ventas.

Palabra clave

Orienta tu herramienta de IA en la dirección correcta, ya sea extrayendo información de los datos o creando imágenes usando palabras claves específicas.

“Espero un documento estructurado con recomendaciones prácticas implementables en 6 meses. Utiliza el formato APA.”

04

3. Criterios de Calidad

Tutorial de Uso

¿Qué es?
Los criterios de calidad son estándares o requisitos que deseas que cumpla la respuesta de la IA.
Seleccionar criterios ayuda a obtener resultados más útiles, claros y alineados con tus necesidades.

Ejemplo:

“Incluir ejemplos ilustrativos y una estructura lógica clara.”

Opciones:

- Especificidad y claridad
- Estructura lógica
- Parámetros y restricciones
- Ejemplos ilustrativos
- Instrucciones paso a paso
- Conocimiento previo y suposiciones
- Iteración y refinamiento
- Consideraciones éticas y sesgos

Ejemplo de uso:
Selecciona “Ejemplos ilustrativos” para que la IA incluya casos concretos.

05

4. Formato y Estilo

a) Formato del Contenido

¿Qué es?
El formato del contenido define el tipo de documento o presentación que deseas recibir.
Esto puede ser un blog, un artículo, un guion, un análisis, un email, etc.

Opciones:

Blog, Artículo, Guion, Análisis, Email, Resumen ejecutivo

Ejemplo:

“Resumen ejecutivo” para un informe breve y profesional.

b) Vocabulario

¿Qué es?
El vocabulario determina el nivel de tecnicismo, formalidad o especialización del lenguaje que debe usar la IA.
Esto es útil para adaptar la respuesta a diferentes públicos.

Opciones:

Simple, Complejo, Técnico, Coloquial, Profesional

Ejemplo:

“Técnico” para un público especializado.

c) Tono

¿Qué es?
El tono es la actitud o emoción predominante en el texto generado.
Puede ser formal, informal, persuasivo, serio, humorístico, empático, etc.

Opciones:

Formal, Informal, Persuasivo, Serio, Humorístico, Empático

Ejemplo:

“Redacta la respuesta con un tono empático y alentador.”

Al hablar en público, el "tono" se refiere a la forma en que modulamos nuestra voz para expresar emociones y transmitir mensajes con mayor impacto. Se pueden utilizar diferentes tonalidades para lograr diferentes efectos, desde formal hasta informal, pasando por serio, humorístico o incluso irónico.

Tonalidades comunes al hablar en público:

Formal:

Se caracteriza por un lenguaje preciso y respetuoso, utilizado en situaciones oficiales o ante audiencias importantes.

Informal:

Es más cercano y relajado, ideal para conversaciones casuales o para conectar con el público de manera personal.

Profesional:

Combina la formalidad con un tono más cercano, adecuado para presentaciones empresariales o conferencias.

Autoritario:

Se usa para transmitir seguridad y liderazgo, ideal para situaciones donde se necesita dar instrucciones o tomar decisiones.

Amable y empático:

Se utiliza para generar confianza y cercanía con la audiencia, ideal para situaciones donde se necesita conectar emocionalmente.

Humorístico:

Se usa para hacer que la presentación sea más divertida y atractiva, pero debe usarse con moderación y cuidado.

Serio y reflexivo:

Se usa para transmitir la importancia de un tema y hacer que la audiencia se tome las palabras en serio.

Activo y motivador:

Se usa para energizar a la audiencia y hacerla participar en la presentación.

Solemne:

Se usa en ocasiones especiales o importantes para expresar respeto y solemnidad.

Irónico:

Se usa para generar humor o resaltar la contradicción en una situación, pero debe usarse con cuidado para evitar confusiones.

Otros factores que influyen en el tono:

Velocidad:

Hablar rápido puede indicar nerviosismo o urgencia, mientras que hablar lento puede transmitir calma y reflexión.

Volumen:

Hablar demasiado alto puede resultar agresivo, mientras que hablar demasiado bajo puede ser difícil de escuchar.

Ritmo:

Un ritmo constante y suave puede facilitar la comprensión, mientras que un ritmo irregular puede generar confusión.

Pausas:

Las pausas estratégicas pueden ayudar a enfatizar puntos importantes y a dar tiempo a la audiencia para procesar la información.

Modulación:

Cambiar la altura de la voz puede ayudar a enfatizar palabras o frases clave y a mantener la atención de la audiencia.

Timbre:

La calidad de la voz (claridad, brillo, fuerza) también influye en el tono y en la forma en que la audiencia percibe la comunicación.

d) Estructura de Oraciones

¿Qué es?
Indica cómo deben estar construidas las oraciones: si deben ser cortas y directas, largas y elaboradas, o neutras.
Esto afecta la claridad y el ritmo del texto.

Opciones:

Neutro, Cortas y directas, Largas y elaboradas

Ejemplo:

“Prefiero oraciones cortas y directas para facilitar la lectura.”

e) Nivel de Detalle

¿Qué es?

Define cuánta información y profundidad debe tener la respuesta.

Puede ser conciso, descriptivo, narrativo o muy detallado.

Opciones:

Conciso, Descriptivo, Narrativo, Muy detallado

Ejemplo:

“Necesito una explicación muy detallada, paso a paso para lograr una explicación extensa.”

f) Perspectiva

¿Qué es?
La perspectiva es el punto de vista desde el cual se redacta el texto: primera persona (“yo”), segunda persona (“tú”), tercera persona (“él/ella/ellos”) o impersonal.
Esto influye en la cercanía y el estilo de la comunicación.

Opciones:

Primera persona, Segunda persona, Tercera persona, Impersonal

Ejemplo:

- “Primera persona” para respuestas como “Yo recomiendo...”.
- “Redacta en tercera persona para mantener objetividad.”

06

5. Audiencia y Configuraciones

a) Audiencia

¿Qué es?
A quién va dirigido el contenido.

Opciones:

Estudiantes, Docentes, Clientes, Investigadores, Gerencia, Público General

Ejemplo:

"El contenido está dirigido a profesores universitarios y a estudiantes de nivel graduado."

b) Formato de Salida

¿Qué es?
Es la forma en que deseas recibir la información: tabla, viñetas, párrafos, esquema, gráfica, preguntas y respuestas, etc.
Esto ayuda a organizar la información de manera más útil para ti.

Opciones:

Tabla, Viñetas, Párrafos, Esquema, Gráfica, Q&A

Ejemplo:

“Presenta la información en formato de tabla comparativa.”

c) Complejidad

¿Qué es?
Indica el nivel de dificultad o profundidad del contenido: básico, intermedio o avanzado.
Esto permite adaptar la respuesta al nivel de conocimiento del público objetivo.

Opciones:

Básico, Intermedio, Avanzado

Ejemplo:

“Elabora una explicación avanzada para especialistas en el tema.”

d) Idioma

¿Qué es?

Selecciona el idioma en el que deseas recibir la respuesta: español, inglés o bilingüe.

Esto es útil para audiencias internacionales o bilingües.

Opciones:

Español, Inglés, Bilingüe

Ejemplo:

“Bilingüe” para recibir el contenido en ambos idiomas.

e) Interacción

¿Qué es?

Define si la respuesta será un solo mensaje (respuesta única) o un diálogo interactivo (con posibilidad de preguntas y respuestas continuas).

Esto es útil para simulaciones de conversación o asesoría continua.

Opciones:

Respuesta única, Conversación interactiva

Ejemplo:

“Prefiero una conversación interactiva para poder hacer preguntas de seguimiento.”

f) Prompt “negativo” (Qué evitar)

¿Qué es?
Aquí puedes especificar temas, enfoques, palabras o elementos que NO deseas que aparezcan en la respuesta.
Esto ayuda a evitar información irrelevante o no deseada.

Ejemplo:

“Evitar tecnicismos avanzados”, “Evita utilizar fuentes no confiables, no verificables y que no hayan sido auditadas por pares profesionales”.

07

6. Parámetros Avanzados (Controladores de la muestra)

Los LLM no predicen formalmente un solo token. Más bien, los LLM predicen probabilidades de lo que el siguiente token podría ser, con cada token en el vocabulario de LLM obteniendo una probabilidad. Estas probabilidades en los tokens se muestrean para determinar cuál será el próximo token producido.

Temperatura, “Top-K” y “Top-P” son los ajustes de configuración más comunes que determinan cómo se procesan las probabilidades de los tokens predeterminado para elegir un único token de salida. (Lee Boonstra, 2025)

¿Qué es?

La “temperatura” en los prompts de inteligencia artificial es un parámetro que controla la creatividad y el grado de incertidumbre en las respuestas generadas. Funciona ajustando la probabilidad con la que el modelo selecciona palabras o frases menos comunes. En esencia, una temperatura más baja produce respuestas más predecibles y estructuradas, mientras que una temperatura más alta genera resultados más variados y creativos, pero también potencialmente menos coherentes.

a) Temperatura

Por ejemplo:

Temperatura baja (0.2):

Si se pregunta: “Describe un atardecer en la playa”, una respuesta con temperatura baja podría ser:

“El atardecer en la playa es un espectáculo de colores cálidos, donde el sol se hunde lentamente en el horizonte, tiñendo el cielo de tonos naranjas y rosados.”

Este enfoque es directo, preciso y menos propenso a desviarse del tema.

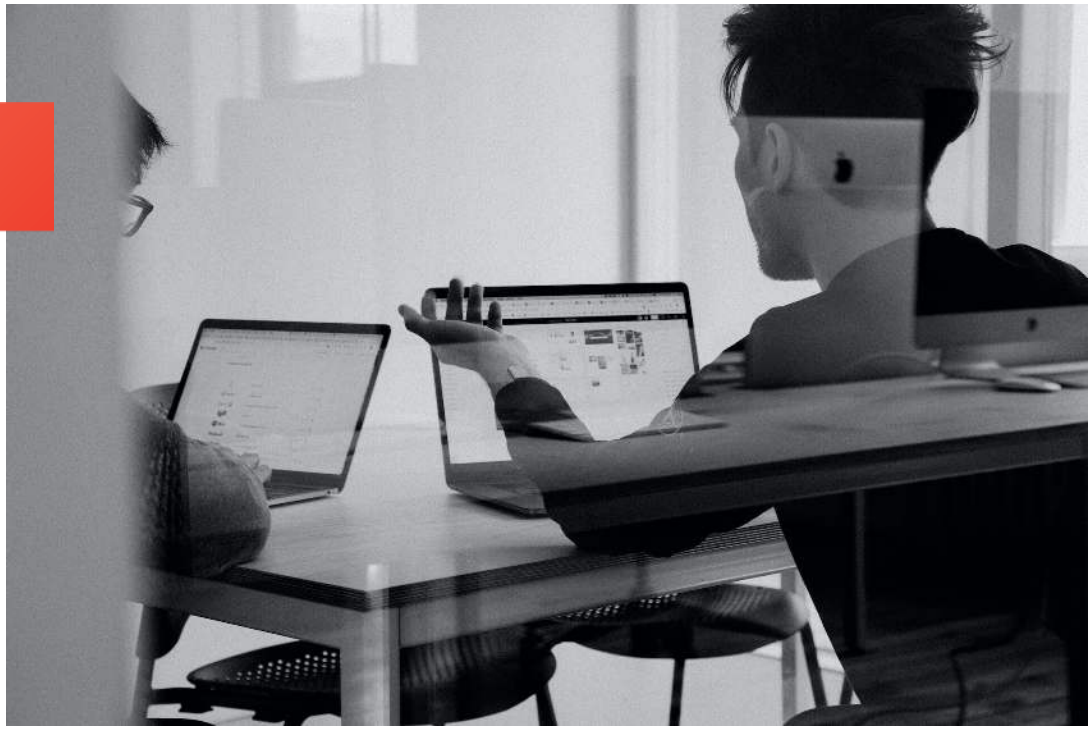
Temperatura alta (0.8):

La misma pregunta podría generar algo como:

“El atardecer en la playa es un lienzo en movimiento, un susurro dorado que se mezcla con el murmullo de las olas, mientras el sol desaparece como un sueño en el horizonte.”

Aquí, la respuesta es más poética y creativa, pero podría no ajustarse tan bien si se busca algo técnico o factual.

Usar la temperatura permite personalizar el tono y estilo de las respuestas según las necesidades del usuario y el tipo de audiencia al que se dirige.



Controla la creatividad de la IA (0 = determinista, 2 = muy creativo).

Ejemplo:

Usa 0.7 para respuestas balanceadas.

b) Top p

Este método controla la creatividad y diversidad de las respuestas de la IA de una manera diferente al ajuste de temperatura. En lugar de modificar la probabilidad de seleccionar palabras individuales, “Top P” establece un límite acumulativo de probabilidad. Esto significa que la IA considerará únicamente las palabras más probables hasta que el porcentaje acumulado alcance el valor especificado por “Top P”.

Por ejemplo:

Top p = 1: IA evaluará todas las palabras posibles, lo que permite una mayor diversidad y creatividad. Esto es útil cuando se busca explorar ideas o recibir respuestas menos restringidas.

Top p = 0.5: La IA limitará sus opciones a las palabras más probables que sumen un 50 % de la probabilidad total, descartando opciones menos relevantes o inusuales. Esto resulta en respuestas más directas y predecibles.

Ejemplo práctico:

Supongamos que le pedimos a la IA una descripción de un bosque al amanecer:

Top p = 1:

“El bosque despierta con un suave resplandor dorado, los árboles proyectan sombras alargadas mientras el canto de los pájaros llena el aire con promesas de un nuevo día.”

Top p = 0.5:

“El bosque amanece con luz dorada y los pájaros cantan.”

Como se observa, un valor más bajo de *top p* tiende a simplificar las respuestas, mientras que un valor más alto ofrece mayor riqueza y detalle, aunque puede ser menos específico dependiendo del contexto deseado.

¿Qué es?
Limita la probabilidad acumulada de las palabras generadas (1 = sin límite, 0.5 = más restrictivo).

Ejemplo:

Usa 1 para máxima variedad, 0.5 para respuestas más predecibles.

c) Top k

El muestreo Top-K selecciona los K tokens más probables de los tokens predichos por la distribución del modelo. Cuanto más alto sea el “Top-K”, más creativo y variado será el resultado del modelo; el más bajo top-K, más fáctico es el resultado del modelo. Un K-top de 1 es equivalente a mayor “aleatoriedad” en la selección de los próximos tokens.

08

7. Mejores Prácticas

Tutorial de Uso

- Sé específico y claro.
- Proporciona contexto relevante.
- Define el formato y la audiencia.
- Especifica limitaciones y exclusiones.
- Incluye ejemplos cuando sea posible.
- Considera aspectos éticos.

Tutorial de Uso

"Tutorial de Uso" es una guía esencial para maximizar la eficacia de la comunicación con inteligencia artificial en el ámbito empresarial. A través de técnicas de prompting y ejemplos prácticos, este libro enseña cómo formular preguntas que optimizan resultados y reducen costos operativos. Con un enfoque en la claridad y la especificidad, este recurso es ideal para quienes buscan potenciar la interacción con modelos de lenguaje avanzados.