Temario curso Inteligencia artificial y ChatGPT

Índice

[Objetivo del curso 2](#_Toc162860453)

[Inteligencia Artificial, conceptos fundamentales. 2](#_Toc162860454)

[¿Qué es la inteligencia? 2](#_Toc162860455)

[¿Qué es la inteligencia artificial? 2](#_Toc162860456)

[Diferencias entre Inteligencia Artificial e Inteligencia Natural 3](#_Toc162860457)

[Origen: 3](#_Toc162860458)

[Naturaleza: 3](#_Toc162860459)

[Historia de la inteligencia artificial. Inicio y actualidad. 3](#_Toc162860460)

[Década de 1950 3](#_Toc162860461)

[Década de 1960-1970 4](#_Toc162860462)

[Década 1980 4](#_Toc162860463)

[Década 1990 5](#_Toc162860464)

[Década 2000 5](#_Toc162860465)

[Década 2010 5](#_Toc162860466)

[Actualidad 6](#_Toc162860467)

[Machine Learning y Redes Neuronales 7](#_Toc162860468)

[¿Qué es y cómo funciona el Machine Learning? 7](#_Toc162860469)

[¿Qué son y cómo funcionan las Redes Neuronales? 7](#_Toc162860470)

[ChatGPT. Origen, Evolución y Diferencia entre versiones 9](#_Toc162860471)

[Todas sus versiones 9](#_Toc162860472)

[Prompt definición y conceptos 9](#_Toc162860473)

[¿Qué es un Prompt? 9](#_Toc162860474)

[Componentes de un Prompt 10](#_Toc162860475)

[Aplicación de Prompt en tareas especificas 10](#_Toc162860476)

[Diseño Efectivo de Prompts: Creación de Instrucciones claras y precisas. 11](#_Toc162860477)

[¿Cómo dar instrucciones claras y precisas? 11](#_Toc162860478)

[Ejemplos y mejores practicas 12](#_Toc162860479)

[Optimización y Ajuste de Prompts en Ingeniería de Prompt 12](#_Toc162860480)

[Pasos para la optimización de un Prompt 12](#_Toc162860481)

[Formatos de salida 13](#_Toc162860482)

[Evaluación de Prompts 14](#_Toc162860483)

[Estrategias para mantener conversaciones con ChatGPT 14](#_Toc162860484)

[¿En qué casos usar ChatGPT? 15](#_Toc162860485)

# Objetivo del curso

Conocer la historia de la inteligencia artificial, como funciona, como fue creada y aprender a comunicarte de manera correcta con ChatGPT.

# Inteligencia Artificial, conceptos fundamentales.

La Inteligencia Artificial (IA) es un campo fascinante que ha capturado la imaginación de científicos, ingenieros y la sociedad en general. En el contexto de la ingeniería de prompts, comprender qué es la IA y cómo se relaciona con la inteligencia natural es fundamental para desarrollar sistemas de conversación eficaces y comprender su potencial y sus limitaciones.

## ¿Qué es la inteligencia?

La inteligencia se ha definido de muchas maneras, incluyendo: la capacidad de lógica, comprensión, autoconciencia, aprendizaje, conocimiento emocional, razonamiento, planificación, creatividad, pensamiento crítico y resolución de problemas. En términos más generales, se puede describir como la capacidad de percibir o inferir información, y retenerla como conocimiento para aplicarlo a comportamientos adaptativos dentro de un entorno o contexto.

Algunas definiciones de investigadores de la psicología y el aprendizaje:

* **Ch. Spearman:** Capacidad unitaria para resolver problemas y crear nuevos contenidos.
* **L. G. Humphreys**: Conjunto de habilidades para adaptarse al entorno.
* **L. S. Gottfredson**: Capacidad mental general que incluye la habilidad de razonar, planificar, resolver problemas, pensar en abstracto, comprender ideas complejas, aprender rápido y aprender de la experiencia, que es más que una destreza académica o del aprendizaje por medio de libros.

## ¿Qué es la inteligencia artificial?

La Inteligencia Artificial se refiere a la capacidad de las máquinas para realizar tareas que normalmente requieren inteligencia humana. Esto incluye el aprendizaje, el razonamiento, la percepción, la comprensión del lenguaje natural y la toma de decisiones. En esencia, la IA busca imitar las capacidades cognitivas humanas utilizando algoritmos y datos.

1. Aprendizaje Automático (Machine Learning): Es una subdisciplina de la IA que se centra en el desarrollo de algoritmos que permiten a las máquinas aprender patrones y tomar decisiones basadas en datos sin una programación explícita.
2. Redes Neuronales Artificiales: Inspiradas en la estructura del cerebro humano, las redes neuronales artificiales son modelos computacionales que pueden aprender y realizar tareas complejas como reconocimiento de patrones y procesamiento de lenguaje natural.
3. Procesamiento del Lenguaje Natural (NLP): Se refiere a la capacidad de las computadoras para entender, interpretar y generar lenguaje humano de manera similar a los seres humanos.
4. Visión por Computadora: Implica el desarrollo de sistemas que pueden interpretar y comprender el contenido visual, como imágenes y videos.

## Diferencias entre Inteligencia Artificial e Inteligencia Natural

La principal diferencia entre la inteligencia artificial y la inteligencia natural radica en su origen y naturaleza.

### Origen:

Inteligencia Artificial: Es creada por humanos a través del diseño y la programación de sistemas informáticos.

Inteligencia Natural: Es inherente a los seres vivos, especialmente a los humanos, y se desarrolla a través de procesos biológicos y experiencia.

### Naturaleza:

Inteligencia Artificial: Se basa en algoritmos y datos para imitar el comportamiento humano en tareas específicas. A menudo es especializada y puede carecer de comprensión profunda o conciencia.

Inteligencia Natural: Es inherentemente compleja y abarca una amplia gama de habilidades cognitivas, emocionales y sociales. Se desarrolla a lo largo del tiempo a través de la interacción con el entorno y la experiencia.

# Historia de la inteligencia artificial. Inicio y actualidad.

## Década de 1950

Alan Turing, pionero en ciencias de la computación, propuso en **1950** una prueba para determinar si una máquina puede exhibir un comportamiento inteligente equivalente al de un ser humano. Esta prueba, conocida como la Prueba de Turing, sigue siendo un concepto fundamental en el campo de la IA y ha estimulado debates sobre la naturaleza de la inteligencia y la capacidad de las máquinas para replicarla.

En **1952**, Arthur Samuel creó un software capaz de aprender a jugar al ajedrez de forma autónoma.

El término "inteligencia artificial" fue acuñado por John McCarthy en **1956** durante una conferencia en Dartmouth College. Este evento marcó el inicio oficial del campo de la IA y reunió a destacados científicos de diversas disciplinas para explorar el potencial de las máquinas para exhibir comportamientos inteligentes.

En **1959**, Arthur Samuel acuñó el término Machine Learning mientras trabajaba en IBM. Por su parte, John McCarthy y Marvin Minsky fundaron el MIT Artificial Intelligence Project.

## Década de 1960-1970

En **1963**, John McCarthy también creó el «AI Lab» en la Universidad de Stanford.

En **1966**, el informe estadounidense ALPAC puso de manifiesto la falta de avances en la investigación de la traducción automática destinada a traducir simultáneamente la lengua rusa en el contexto de la Guerra Fría.

Muchos proyectos financiados por el gobierno estadounidense fueron cancelados.

Del mismo modo, en **1973**, el gobierno británico publicó su informe Lighthill en el que destacaba las decepciones de la investigación en IA. Una vez más, los proyectos de investigación fueron reducidos por los recortes presupuestarios. Este periodo de duda duró hasta 1980, y ahora se denomina el **primer invierno(crisis) de la IA**.

## Década 1980

La década de **1980** presenció un auge en la investigación y aplicación de la IA en áreas como el procesamiento del lenguaje natural, la visión por computadora y los sistemas expertos. Los sistemas expertos, en particular, se convirtieron en una aplicación popular de la IA en campos como la medicina, la ingeniería y la gestión empresarial.

En **1985** se desarrolla el sistema DENDRAL, que utiliza técnicas de inteligencia artificial para identificar estructuras moleculares desconocidas.

En **1987** Martin Fischles y Oscar Firschein describieron los atributos de un agente inteligente. Al intentar describir con un mayor ámbito (no solo la comunicación) los atributos de un agente inteligente, la IA se ha expandido a muchas áreas que han creado ramas de investigación enormes y diferenciadas. Dichos atributos del agente inteligente son:

1. Tiene actitudes mentales tales como creencias e intenciones.
2. Tiene la capacidad de obtener conocimiento, es decir, aprender.
3. Puede resolver problemas, incluso descomponiendo problemas complejos en otros más simples.
4. Capaz de realizar operaciones más complejas.
5. Entiende.
6. Planifica.
7. Conoce los límites de sus propias habilidades y conocimientos.
8. Puede distinguir a pesar de la similitud de las situaciones.
9. Puede ser original, creando incluso nuevos conceptos o ideas, y hasta utilizando analogías.
10. Puede generalizar.
11. Puede percibir y modelar el mundo exterior.
12. Puede entender y utilizar el lenguaje y sus símbolos.

Este fue el **segundo invierno(crisis) de la IA**. Las empresas perdieron el interés por los sistemas expertos. Los gobiernos de Estados Unidos y Japón abandonaron sus proyectos de investigación y se gastaron miles de millones de dólares para nada.

## Década 1990

Durante la década de 1990, el enfoque en el aprendizaje automático y las redes neuronales resurgió con el desarrollo de algoritmos como las máquinas de vectores de soporte (SVM) y las redes neuronales convolucionales (CNN). Estos avances sentaron las bases para una nueva ola de aplicaciones de IA en áreas como la minería de datos, la detección de fraudes y el reconocimiento de patrones.

En **1997**, la historia de la IA estuvo marcada por un acontecimiento importante. La IA Deep Blue de IBM triunfó sobre el campeón mundial de ajedrez Garry Kasparov. Por primera vez, el hombre fue derrotado por la máquina.

## Década 2000

En la década de 2000, el surgimiento del concepto de "Big Data" impulsó el desarrollo de técnicas de IA capaces de manejar y procesar grandes volúmenes de datos. Además, se produjeron avances significativos en el campo del aprendizaje profundo, una rama del aprendizaje automático que utiliza redes neuronales profundas para aprender representaciones de datos de alto nivel. Estos avances permitieron a los sistemas de IA realizar tareas más complejas y sofisticadas, como el reconocimiento de imágenes, la traducción automática y la conducción autónoma.

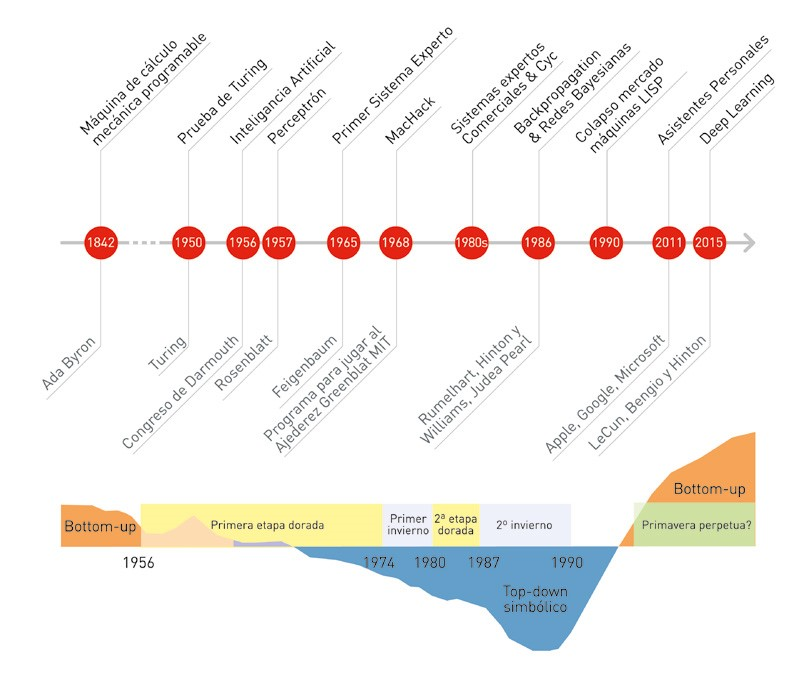
## Década 2010

La década de 2010 fue testigo de un rápido avance en la adopción de la inteligencia artificial en una variedad de aplicaciones prácticas. Los algoritmos de IA se utilizaron cada vez más en áreas como la atención médica, la banca, el comercio electrónico y la publicidad en línea. Además, hubo avances significativos en el desarrollo de asistentes virtuales y sistemas de reconocimiento de voz, como Siri de Apple, Alexa de Amazon y Google Assistant, que se convirtieron en parte integral de la vida cotidiana de millones de personas en todo el mundo.

En **2012**, Andrew Ng alimentó una red neuronal con 10 millones de vídeos de YouTube como serie de datos de entrenamiento. Gracias al Deep Learning, esta red neuronal aprendió a reconocer un gato sin que se le enseñara lo que es un gato.

En **2014** un ordenador ha logrado superar con éxito el test de Turing, haciendo creer a un interrogador que es una persona quien responde sus preguntas en un certamen organizado en Londres por la Universidad de Reading

En **2016** un ordenador de Google venció al campeón mundial del juego milenario “Go”. Un programa informático desarrollado por la compañía británica Google DeepMind consiguió vencer, por primera vez, a un campeón profesional del milenario juego de origen oriental Go.



## Actualidad

La década de 2020 ha sido testigo de una mayor integración de la inteligencia artificial en la sociedad, impulsada por la transformación digital en diversos sectores. Los avances en el procesamiento del lenguaje natural, la visión por computadora y el aprendizaje automático han llevado a la creación de sistemas de IA cada vez más sofisticados y autónomos. Sin embargo, también ha habido un mayor énfasis en la ética y la responsabilidad en el desarrollo y uso de la IA, especialmente en áreas sensibles como la privacidad de los datos, el sesgo algorítmico y la toma de decisiones automatizada.

En la actualidad, la inteligencia artificial está presente en todos los aspectos de la vida, desde la atención médica y la educación hasta el entretenimiento y la seguridad. Los sistemas de IA alimentan plataformas de redes sociales, recomendaciones de contenido en línea, vehículos autónomos, sistemas de detección de fraudes y mucho más. A medida que la tecnología continúa avanzando, es probable que la inteligencia artificial siga desempeñando un papel cada vez más importante en la sociedad, con implicaciones tanto positivas como desafiantes que deben abordarse de manera proactiva y ética.

# Machine Learning y Redes Neuronales

## ¿Qué es y cómo funciona el Machine Learning?

El Machine Learning es una rama de la inteligencia artificial que se enfoca en el desarrollo de algoritmos y modelos que permiten a las computadoras aprender de los datos y tomar decisiones sin ser programadas explícitamente para realizar una tarea específica. En lugar de seguir reglas predefinidas, los algoritmos de Machine Learning utilizan datos para aprender patrones y tomar decisiones informadas.

El proceso básico de Machine Learning consta de varias etapas:

1. **Adquisición de Datos**: En esta etapa, se recopilan y preparan los datos necesarios para el entrenamiento del modelo. Estos datos pueden provenir de diversas fuentes, como bases de datos, sensores, archivos CSV, etc.
2. **Preprocesamiento de Datos**: Los datos se preprocesan para eliminar ruido, manejar valores faltantes y convertirlos en un formato adecuado para el modelo de Machine Learning.
3. **Selección de Características**: Se seleccionan las características relevantes que se utilizarán para entrenar el modelo. Esto puede implicar la extracción de características útiles de los datos brutos.
4. **Entrenamiento del Modelo**: En esta etapa, el modelo de Machine Learning se entrena utilizando un algoritmo específico y los datos de entrenamiento. Durante el entrenamiento, el modelo ajusta sus parámetros para minimizar el error en la predicción.
5. **Validación y Evaluación**: Una vez que el modelo está entrenado, se evalúa su rendimiento utilizando datos de validación o prueba. Se calculan métricas de rendimiento para determinar la precisión y la eficacia del modelo.
6. **Ajuste del Modelo**: Si es necesario, se realizan ajustes adicionales al modelo para mejorar su rendimiento, como la optimización de hiperparámetros o la selección de un algoritmo diferente.
7. **Despliegue y Uso**: Finalmente, el modelo entrenado se despliega en un entorno de producción y se utiliza para hacer predicciones o tomar decisiones en tiempo real.

## ¿Qué son y cómo funcionan las Redes Neuronales?

Las redes neuronales son un conjunto de algoritmos y modelos inspirados en la estructura y funcionamiento del cerebro humano. Están diseñadas para reconocer patrones y procesar información de manera similar a como lo hacen las neuronas en el cerebro. Se componen de capas de nodos interconectados, cada uno de los cuales realiza operaciones matemáticas en los datos de entrada y transfiere la información a través de la red.

Elementos clave de las redes neuronales:

* Neuronas (nodos): Son unidades básicas que reciben información, realizan cálculos y transmiten resultados a nodos posteriores. Cada neurona está asociada con una función de activación que determina su salida en función de la información que recibe.
* Pesos y conexiones: Las conexiones entre las neuronas tienen pesos asignados. Estos pesos representan la importancia o la fuerza de la conexión entre dos nodos y se ajustan durante el entrenamiento de la red para mejorar su capacidad predictiva.
* CAPAS: Las redes neuronales se componen de múltiples capas:
  + Capa de entrada: Recibe datos de entrada.
  + Capas ocultas (intermedias): Realizan cálculos complejos para procesar la información.
  + Capa de salida: Produce los resultados finales.

El funcionamiento de una red neuronal se basa en la propagación hacia adelante (forward propagation) y la propagación hacia atrás (backpropagation):

* Propagación hacia Adelante: Durante esta etapa, los datos de entrada se introducen en la red neuronal a través de la capa de entrada. Los datos se propagan a través de las capas ocultas, donde se realizan cálculos ponderados y se aplican funciones de activación para generar una salida.
* Backpropagation: Una vez que se genera la salida, se compara con la salida deseada para calcular el error de predicción. Este error se propaga hacia atrás a través de la red, ajustando los pesos de las conexiones para minimizar el error.

El entrenamiento de una red neuronal implica los siguientes pasos:

* Inicialización de pesos: Se asignan valores aleatorios a los pesos y sesgos de las conexiones entre las neuronas.
* Propagación hacia adelante (Forward Propagation): Los datos de entrada se pasan a través de la red neuronal, capa por capa, y se calculan las salidas de cada neurona utilizando una función de activación.
* Cálculo de la función de pérdida (Loss Function): Se compara la salida predicha de la red con las respuestas reales, calculando una medida de la discrepancia entre ellas (pérdida o error).
* Propagación hacia atrás (Backward Propagation): Se utiliza un algoritmo llamado descenso de gradiente para calcular la derivada de la función de pérdida con respecto a cada peso y sesgo en la red. Estas derivadas indican cómo cambiar los pesos y sesgos para reducir la pérdida.
* Actualización de pesos (Weight Update): Se ajustan los pesos y sesgos en la dirección que reduce la pérdida utilizando un método de optimización, como el descenso de gradiente estocástico (SGD) o algoritmos más avanzados como Adam o RMSprop.
* Repetición del proceso: Se repiten los pasos 2 a 5 para cada lote de datos de entrenamiento durante múltiples épocas (iteraciones completas a través de todos los datos de entrenamiento) hasta que la red neuronal converge y la pérdida se minimiza.

# ChatGPT. Origen, Evolución y Diferencia entre versiones

El origen de ChatGPT se remonta a los avances en inteligencia artificial, específicamente en el campo del procesamiento del lenguaje natural (PLN) y los modelos de lenguaje generativo. En junio de 2020, OpenAI lanzó la primera versión de ChatGPT, conocida como GPT-3 (Generative Pre-trained Transformer 3), que marcó un hito significativo en el desarrollo de sistemas de conversación basados en IA.

GPT-3 fue diseñado utilizando la arquitectura Transformer, que utiliza una atención auto-referencial y capas de auto-atención para modelar secuencias de texto. Con 175 mil millones de parámetros, GPT-3 fue el modelo de lenguaje más grande y avanzado hasta ese momento, capaz de generar respuestas coherentes y contextualmente relevantes en una variedad de situaciones.

## Todas sus versiones

1. **GPT** (Generative Pre-trained Transformer): Esta fue la primera versión de la serie GPT, lanzada en 2018 por OpenAI. Aunque tenía menos parámetros que las versiones posteriores, estableció las bases para el desarrollo de modelos de lenguaje generativo basados en la arquitectura Transformer.
2. **GPT-2:** Lanzado en febrero de 2019, GPT-2 fue una versión mejorada de GPT que contenía 1.5 mil millones de parámetros. Esta versión demostró capacidades sorprendentes en la generación de texto, lo que llevó a OpenAI a restringir inicialmente el acceso al modelo completo debido a preocupaciones sobre su potencial para generar contenido falso y engañoso.
3. **GPT-3**: Lanzado en junio de 2020, GPT-3 representó un avance significativo en el campo de los modelos de lenguaje generativo. Con 175 mil millones de parámetros, GPT-3 fue el modelo de lenguaje más grande y avanzado hasta la fecha de su lanzamiento, capaz de realizar una amplia gama de tareas de PLN con resultados impresionantes.
4. **GPT-4**: lanzamiento de GPT-4 en marzo de 2023, el alejamiento del enfoque de desarrollo de software gratis y abierto fue de la mano con el uso gratuito en condiciones científicamente comprensibles. La situación competitiva de la empresa y las medidas de seguridad se citaron como justificación. El cofundador de OpenAI describió más tarde el enfoque abierto como un error.
5. **ChatGPT Plus**: es un servicio de suscripción para ChatGPT y es propiedad de OpenAI​ La suscripción proporciona acceso al modelo GPT-4 de OpenAI. El servicio brinda a los suscriptores acceso consistente durante los períodos de uso máximo, tiempos de respuesta acelerados y acceso preferencial a nuevas funciones, incluidos GPT-4 y los próximos complementos de ChatGPT.

# Prompt definición y conceptos

## ¿Qué es un Prompt?

En el ámbito de la ingeniería de Prompt, los "Prompts" son instrucciones o guiones que se proporcionan a modelos de lenguaje, como GPT (Generative Pre-trained Transformer), para guiar su generación de texto en una dirección específica. Estas instrucciones pueden variar en complejidad y longitud, desde simples palabras clave hasta párrafos detallados. La eficacia de un prompt radica en su capacidad para comunicar claramente la tarea deseada al modelo de lenguaje.

### Componentes de un Prompt

* Instrucciones Claras: Deben ser lo suficientemente claras para comunicar la tarea deseada al modelo de lenguaje.
* Contexto Relevante: Proporcionar contexto relevante puede ayudar al modelo a comprender mejor la solicitud y generar una respuesta más precisa.
* Palabras Clave: Incluir palabras clave relacionadas con la tarea puede orientar al modelo hacia el tipo de respuesta deseada.
* Restricciones: Pueden incluir restricciones específicas sobre el estilo, tono o longitud de la respuesta.
* Ejemplos: A veces, se pueden incluir ejemplos de respuestas deseadas para ayudar al modelo a comprender mejor la tarea.

## Aplicación de Prompt en tareas especificas

Los prompts se utilizan en una variedad de tareas específicas en el campo de la ingeniería de prompt. Algunos ejemplos incluyen:

1. Generación de Texto Creativo:

Los prompts se utilizan para guiar la generación de textos creativos, como historias, poemas o diálogos. Por ejemplo, un prompt para generar un poema podría ser: "Escribe un poema sobre la belleza de la naturaleza".

1. Traducción Automática:

En tareas de traducción automática, los prompts se utilizan para solicitar la traducción de un texto de un idioma a otro. Por ejemplo, un prompt para traducción podría ser: "Traduce el siguiente párrafo del inglés al español".

1. Resumen de Texto:

Los prompts se utilizan para solicitar un resumen de un texto largo o complejo. Por ejemplo, un prompt para resumir un artículo científico podría ser: "Resume los hallazgos principales de este estudio en no más de 100 palabras".

1. Diálogo y Conversación:

En tareas de generación de diálogos o conversaciones, los prompts se utilizan para establecer el contexto y los participantes de la conversación. Por ejemplo, un prompt para generar un diálogo entre un cliente y un asistente virtual podría ser: "Simula una conversación entre un cliente que necesita ayuda con su factura y un asistente virtual de servicio al cliente".

1. Búsqueda de Respuestas:

Los prompts se utilizan para solicitar respuestas a preguntas específicas. Por ejemplo, un prompt para buscar información sobre un tema podría ser: "¿Cuáles son las causas y consecuencias del cambio climático?".

En resumen, los prompts son herramientas fundamentales en la ingeniería de Prompt, ya que permiten dirigir y personalizar la generación de texto de modelos de lenguaje como GPT para una amplia gama de tareas específicas.

# Diseño Efectivo de Prompts: Creación de Instrucciones claras y precisas.

En el campo de la ingeniería de prompt, el diseño efectivo de prompts es fundamental para obtener resultados precisos y relevantes al utilizar modelos de lenguaje como ChatGPT. Crear prompts claros y precisos implica entender la tarea deseada, comunicarla de manera efectiva al modelo y proporcionar la orientación necesaria para obtener la respuesta deseada.

## ¿Cómo dar instrucciones claras y precisas?

1. **Comprender la tarea:**

Antes de diseñar un prompt, es crucial comprender completamente la tarea que se desea realizar. Esto implica:

* identificar el propósito de la tarea,
* los requisitos específicos
* y el tipo de respuesta esperada.

Por ejemplo, si la tarea implica generar un resumen de un artículo, es importante conocer los puntos clave del artículo y la longitud deseada del resumen.

1. **Identificar las palabras claves:**

Las palabras clave son elementos esenciales en un prompt, ya que ayudan a dirigir la atención del modelo hacia aspectos específicos de la tarea. Estas palabras deben ser **relevantes y específicas** para la tarea en cuestión. Por ejemplo, si la tarea implica escribir un poema sobre el amor, palabras clave como "amor", "pasión", "corazón" pueden ser incluidas en el prompt.

1. **Proporcionar contexto relevante:**

El contexto es fundamental para que el modelo comprenda la tarea y genere una respuesta precisa. Esto puede incluir:

* información adicional sobre el tema,
* el propósito de la tarea
* y cualquier restricción relevante.

Por ejemplo, si la tarea implica traducir un texto, es importante especificar los idiomas de origen y destino, así como el tipo de texto a traducir.

1. **Establecer restricciones:**

Es importante establecer restricciones claras en el prompt para guiar la generación de texto del modelo. Estas restricciones pueden incluir:

* limitaciones en cuanto a la longitud de la respuesta,
* el estilo de redacción
* o incluso la exclusión de ciertos temas.

1. **Establecer un rol:**

Si bien no es 100% necesario al darle un rol especifico en el prompt lo que hará es “ponerse un filtro” a la hora de generar la respuesta correcta. Algunos ejemplos de roles son:

* Asistente laboral
* Asistente de Investigación
* Editor de Texto
* Programador
* Analista de Datos
* Facilitador de tomas de decisiones.

Por ejemplo, si la tarea implica generar un diálogo, se pueden establecer restricciones en cuanto al número de turnos de diálogo o la longitud total del intercambio.

## Ejemplos y mejores practicas

**Prompt para Resumen de Texto:** "Resume los hallazgos principales de este estudio en no más de 100 palabras. El estudio analiza el impacto del cambio climático en la biodiversidad marina."

**Prompt para Generación de Diálogos:** "Simula una conversación entre un cliente y un agente de servicio al cliente para resolver un problema de facturación. El cliente está molesto por un cargo incorrecto en su factura."

**Prompt para Traducción Automática:** "Traduce el siguiente párrafo del español al inglés. El texto trata sobre las nuevas tecnologías en el campo de la medicina."

**Prompt para Generación de Poemas:** "Escribe un poema sobre la belleza de la naturaleza. Incluye imágenes de árboles, flores y ríos."

# Optimización y Ajuste de Prompts en Ingeniería de Prompt

En el contexto de la ingeniería de prompt, la optimización y ajuste de prompts son procesos fundamentales para mejorar la efectividad de los modelos de lenguaje, como ChatGPT. Estos procesos implican la revisión y modificación de los prompts para obtener mejores resultados en diferentes tareas y escenarios.

## Pasos para la optimización de un Prompt

* **Análisis de Resultados:**

Antes de optimizar un prompt, es crucial realizar un análisis exhaustivo de los resultados generados por el modelo en respuesta al prompt original. Esto implica revisar la calidad, relevancia y coherencia de las respuestas para identificar áreas de mejora. Además, es importante considerar la retroalimentación del usuario final para comprender mejor las necesidades y expectativas.

* **Refinamiento de Palabras Clave:**

Una técnica efectiva para mejorar un prompt es refinar las palabras clave utilizadas para guiar al modelo. Esto implica seleccionar términos más específicos y relevantes para la tarea en cuestión, lo que ayuda al modelo a generar respuestas más precisas y centradas. Por ejemplo, en lugar de usar "inteligencia artificial", se puede refinar a "aplicaciones de inteligencia artificial en la atención médica".

* **Incorporación de Ejemplos:**

Incluir ejemplos concretos en el prompt puede ayudar al modelo a comprender mejor la tarea y generar respuestas más precisas. Estos ejemplos pueden proporcionar contextos específicos y guiar al modelo hacia el tipo de respuesta deseada. Por ejemplo, en un prompt para generar un resumen, se pueden incluir ejemplos de resúmenes bien estructurados para ilustrar el formato y contenido esperados.

* **Experimentación Iterativa:**

El proceso de optimización y ajuste de prompts a menudo requiere experimentación iterativa, donde se prueban diferentes variantes del prompt y se evalúan los resultados. Esto puede implicar modificar la estructura del prompt, ajustar las palabras clave o agregar restricciones adicionales. La experimentación iterativa permite identificar la combinación óptima de elementos del prompt que maximizan la calidad de las respuestas del modelo.

* **Convierte tareas complejas en subtareas más simples:**

Para tareas en las que se requieren múltiples conjuntos de instrucciones independientes para manejar diferentes casos, puede ser beneficioso comenzar clasificando el tipo de consulta y utilizar esa clasificación para determinar qué instrucciones se necesitan. Por ejemplo, si vas a hacer el lanzamiento de un producto, es mejor pedirle al ChatGPT que te enumere las tareas y después trabajar sobre ellas.

¡Un pequeño truco al terminar de optimizar tu prompt recomiendo pedirle a ChatGPT que te realice una plantilla autocompletable para llegar a la misma respuesta, sintetizando más aun el trabajo de usar el chat! Esto es muy útil cuando tenes que transformar gran cantidad de texto plano en un nuevo formato.

# Formatos de salida

Parte importante de un Prompt es el formato de salida que le solicitamos a ChatGPT ya que a partir de este se nos hará más fácil trabajar con los datos solicitados, a continuación, verás una lista con algunos ejemplos de formatos de salida que puede darte ChatGPT.

1. Texto plano:

Ejemplo: "La fotosíntesis es el proceso mediante el cual las plantas y otros organismos convierten la luz solar en energía química para alimentarse y crecer."

1. HTML:

Ejemplo: "<p>El cielo nocturno está lleno de estrellas brillantes que forman constelaciones reconocibles, como la Osa Mayor y Orión.</p>"

1. JSON:

Ejemplo: {"respuesta": "¿Sabías que los robots también tienen sueños? Bueno, al menos hasta que se quedan sin batería."}

1. Markdown:

Ejemplo: "La amistad es como una planta; necesita cuidado y atención para crecer y florecer."

1. XML:

Ejemplo: "<respuesta>En un universo paralelo, los gatos son los gobernantes supremos y los humanos son sus leales siervos que les sirven atún fresco todos los días.</respuesta>"

1. CSV:

Ejemplo: respuesta

"Imagina un mundo donde todos los problemas se resuelven con abrazos en lugar de discusiones. Sería un lugar mucho más cálido y amoroso, ¿no crees?"

1. ¡Salida personalizada, cuanto más contexto mejor!

## Evaluación de Prompts

Una vez optimizados los prompts, es crucial evaluar su efectividad en diferentes escenarios y tareas. Esto puede implicar la comparación de las respuestas generadas por el modelo utilizando diferentes variantes de prompts, así como la recopilación de métricas objetivas, como la precisión, coherencia y relevancia de las respuestas. Además, es importante considerar la retroalimentación cualitativa de los usuarios finales para evaluar la utilidad y adecuación de las respuestas generadas.

Supongamos que estamos utilizando ChatGPT para generar respuestas a preguntas sobre salud. Después de optimizar el prompt utilizando técnicas mencionadas previamente, podemos evaluar su efectividad en términos de precisión y relevancia de las respuestas generadas. Podemos comparar las respuestas generadas por el modelo utilizando el prompt original y el prompt optimizado, y recopilar métricas objetivas, como la proporción de respuestas precisas y la coherencia del contenido. Además, podemos solicitar la retroalimentación de profesionales de la salud para evaluar la utilidad y precisión de las respuestas generadas en diferentes escenarios clínicos.

# Estrategias para mantener conversaciones con ChatGPT

En el ámbito de la ingeniería de prompt orientada al uso de modelos de lenguaje como ChatGPT, mantener conversaciones iterativas es una tarea fundamental para mejorar la calidad y fluidez de la interacción entre humanos y máquinas. Una conversación iterativa implica una serie de intercambios entre un usuario y un modelo de lenguaje, donde cada respuesta del modelo se basa en el contexto proporcionado por el usuario en interacciones anteriores.

Una estrategia clave para mantener conversaciones iterativas es **contextualizar adecuadamente la conversación en curso**. Esto implica recordar y utilizar el contexto proporcionado por el usuario en interacciones anteriores para generar respuestas coherentes y relevantes. Por ejemplo, si un usuario está discutiendo sobre películas y menciona su película favorita, el modelo debe recordar esta información y utilizarla en respuestas posteriores para mantener la coherencia en la conversación.

Los tokens de control son una técnica efectiva para guiar la generación de texto del modelo en una dirección específica durante una conversación iterativa. Estos tokens se utilizan para indicar al modelo el estado actual de la conversación y proporcionar instrucciones adicionales sobre cómo generar la siguiente respuesta. Por ejemplo, **un token de control podría indicar al modelo que responda con una pregunta para continuar la conversación** o que genere una respuesta basada en un tema específico.

**La retroalimentación del usuario** juega un papel fundamental en la mejora de la calidad de la conversación iterativa. Al proporcionar retroalimentación sobre la relevancia, coherencia y fluidez de las respuestas del modelo, los usuarios pueden guiar el comportamiento del modelo y ayudar a optimizar la interacción. Además, la retroalimentación también puede utilizarse para corregir errores o malentendidos en la conversación y mejorar la comprensión del modelo sobre las preferencias y necesidades del usuario.

# ¿En qué casos usar ChatGPT?

Buenas situaciones donde se puede utilizar ChatGPT:

* Generando Texto
* Generando Ideas
* Continuando líneas de pensamiento
* Generando Código
* Actuando como un profesional.

Malas situaciones donde utilizar ChatGPT:

* Matemáticas
* Sin datos de ejemplos
* Actualidad (Finales de 2021)
* Preguntas sesgadas.