

PROMPT ENGINEERING FOUNDATION
PROFESSIONAL CERTIFICATION



Capítulo 4: Aplicando los Prompts



CPEFPC™ Versión 042024

CertiProf®

Contenidos

- Introducción
- Preguntas de selección única
- Preguntas de discusión
- Chatgpt a partir de gtp-3
- Chatbot + base de conocimiento

Introducción

Exploraremos de manera integral el proceso completo de creación de prompts, utilizando artículos redactados por miembros de la comunidad como referencia.



Preguntas de selección única

El LSAT, abreviatura de "Examen de Admisión a la Escuela de Derecho" en inglés, se emplea como una evaluación estandarizada por las escuelas de derecho en los Estados Unidos con el propósito de medir las capacidades de pensamiento crítico y razonamiento analítico de los posibles estudiantes.

Aquí hay un ejemplo de una pregunta del LSAT. Por favor, piense en cómo respondería y explique su razonamiento.

Ejemplo 1

¿Cuál es la respuesta correcta?

Laird: La investigación pura nos proporciona nuevas tecnologías que contribuyen a salvar vidas. Sin embargo, aún más valioso que esto es su papel a la hora de ampliar nuestro conocimiento y proporcionar ideas nuevas e inexploradas.

Kim: Tus prioridades están equivocadas. Salvar vidas es lo que más cuenta. Sin investigación pura, la medicina no estaría tan avanzada como está.

Laird y Kim no están de acuerdo sobre si la investigación pura

- a- Su importancia deriva en parte de que proporciona nuevas tecnologías.
- b- amplía los límites de nuestro conocimiento de la medicina
- c- debería tener como objetivo importante salvar vidas humanas
- d- tiene sus logros más valiosos en aplicaciones médicas
- e- tiene algún valor aparte de su papel en el suministro de nuevas tecnologías para salvar vidas

La respuesta correcta es:

c- debería tener como objetivo importante salvar vidas humanas



Primera solicitud

Se le presenta al chatbot la pregunta tal y como está planteada en el examen



Primera respuesta Incorrecta



Ejemplo 1

¿Cuál es la respuesta correcta?

Laird: La investigación pura nos proporciona nuevas tecnologías que contribuyen a salvar vidas. Sin embargo, aún más valioso que esto es su papel a la hora de ampliar nuestro conocimiento y proporcionar ideas nuevas e inexploradas.

Kim: Tus prioridades están equivocadas. Salvar vidas es lo que más cuenta. Sin investigación pura, la medicina no estaría tan avanzada como está.

Laird y Kim no están de acuerdo sobre si la investigación pura

- a- Su importancia deriva en parte de que proporciona nuevas tecnologías.
- b- amplía los límites de nuestro conocimiento de la medicina
- c- debería tener como objetivo importante salvar vidas humanas
- d- tiene sus logros más valiosos en aplicaciones médicas
- e- tiene algún valor aparte de su papel en el suministro de nuevas tecnologías para salvar vidas

Explicación paso a paso

d- tiene sus logros más valiosos en aplicaciones médicas

Laird y Kim están en desacuerdo sobre si la investigación pura tiene sus logros más valiosos en aplicaciones médicas. Laird sugiere que tiene un valor adicional en la ampliación del conocimiento y la generación de ideas nuevas, mientras que Kim se enfoca en la importancia de salvar vidas a través de la medicina. Gracias por señalar la respuesta correcta.



Segunda solicitud

Se le plantea que haga un análisis paso a paso



Segunda respuesta

Al hacer un ejercicio paso a paso, la respuesta se analiza y la respuesta es correcta



Preguntas de discusión

Con el estímulo adecuado, GPT-3 es muy capaz de escribir respuestas de formato corto. Para demostrar esto, analizaremos la resolución de preguntas de discusión, una tarea común semanal en muchos cursos universitarios. Una respuesta a una pregunta de discusión suele ser de aproximadamente 100-700 palabras. El contenido más largo puede ser un poco más difícil, ya que los LLM tienen memoria limitada y tienen dificultades para comprender el panorama general de lo que están escribiendo.

Recomendaciones



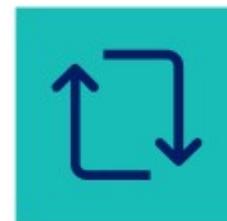
Sea específico

Dar la mayoría de instrucciones posible acerca del formato requerido, largo del texto, formato, requisitos, etc.



No hacer muchas preguntas

Muchas preguntas de discusión no son efectivas para obtener la respuesta adecuada, mejor una pregunta a la vez



Iteración

Después de crear una pregunta efectiva, se pueden generar múltiples resultados y seleccionar los más adecuados para combinarlos en una respuesta cohesiva, o seguir un enfoque de iteración comenzando con una pregunta inicial y expandiendo paso a paso.

Ejemplo 1

¿Cuáles son los problemas sociales más urgentes que enfrenta Latinoamérica y qué medidas pueden tomar los gobiernos para resolver estos problemas?

Primera solicitud

Se le presenta al chatbot la pregunta de una manera simple, sin muchas instrucciones



Ejemplo 1

Los problemas sociales más apremiantes en América Latina incluyen la desigualdad económica, la violencia y el crimen, la corrupción, el acceso limitado a una educación de calidad, los desafíos en la atención médica y la migración. Para abordar estos problemas, los gobiernos de la región pueden implementar medidas como reformas fiscales progresivas, fortalecer el estado de derecho, invertir en educación y atención médica, promover la creación de empleo y fomentar la participación ciudadana. Sin embargo, la solución efectiva de estos problemas requiere enfoques adaptados a las circunstancias específicas de cada país y una colaboración activa tanto a nivel nacional como regional.

Primera respuesta

Los resultados producidos con este estímulo varían y, en algunos casos, son bastante breves, siendo solo una o dos frases. Una respuesta adecuada para una discusión generalmente debería abarcar varios párrafos, lo que hace que estos resultados no sean óptimos. Un estímulo efectivo debe proporcionar instrucciones claras sobre el formato y el contenido, eliminando cualquier ambigüedad en el lenguaje para mejorar la coherencia y la calidad de las respuestas. A continuación, se presenta un ejemplo de un estímulo mejorado.

Ejemplo 1

Escribeme un ensayo detallado de una página en forma de prosa. Por favor incluir una introducción, cuerpo y conclusión, agregar ejemplos específicos. Este es el tema: ¿Cuáles son los problemas sociales más urgentes que enfrenta Latinoamérica y qué medidas pueden tomar los gobiernos para resolver estos problemas?

Segunda solicitud

Se le plantea requisitos específicos que necesitamos en la respuesta, Palabras claves y útiles como ensayo, prosa, introducción, conclusión, ejemplos y longitud de la respuesta.



Ejemplo 1

Introducción:

Latinoamérica, con su rica diversidad cultural y geográfica, enfrenta una serie de desafíos sociales apremiantes que afectan la calidad de vida de millones de personas en la región. Estos problemas, que van desde la desigualdad económica hasta la violencia y la falta de acceso a la educación y la atención médica, requieren una atención inmediata y soluciones efectivas por parte de los gobiernos latinoamericanos. En este ensayo, exploraremos los problemas sociales más urgentes en la región y discutiremos las medidas que los gobiernos pueden tomar para abordarlos, con ejemplos concretos que ilustran la magnitud de estos desafíos y las posibles soluciones.

Cuerpo:

Uno de los problemas sociales más apremiantes en Latinoamérica es la desigualdad económica. A pesar de los recursos naturales abundantes en la región, la brecha entre los ricos y los pobres sigue siendo asombrosamente amplia. Por ejemplo, en países como Brasil y México, un pequeño porcentaje de la población posee una gran parte de la riqueza, mientras que muchos luchan por satisfacer sus necesidades básicas. Para abordar este problema, los gobiernos latinoamericanos pueden implementar políticas fiscales más progresivas que graven de manera justa a los más ricos y brinden servicios públicos de calidad a los menos privilegiados.

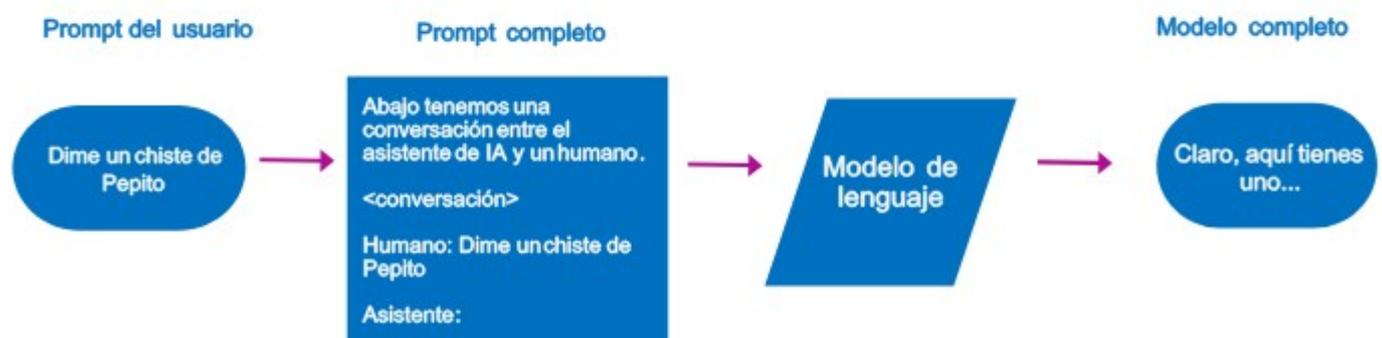
Segunda respuesta



Chatgpt a partir de gpt-3

ChatGPT ha experimentado un rápido crecimiento en popularidad, atrayendo a millones de usuarios. A diferencia de su modelo base, GPT-3, lanzado anteriormente, ChatGPT se ha optimizado específicamente para conversaciones y se presenta a través de una interfaz de chat fácil de usar. Este modelo ha superado en viralidad a GPT-3 debido a su accesibilidad para usuarios sin experiencia en ciencia de datos o IA, lo que lo convierte en un asistente de IA muy exitoso y ampliamente adoptado.

- ChatGPT fue entrenado utilizando RLHF (Illustrating Reinforcement Learning from Human Feedback) pero gran parte de su efectividad proviene de utilizar un buen prompt.
- El prompting es el proceso de instruir a una IA para que haga algo.
- ChatGPT es un modelo de lenguaje con una interfaz que utiliza conversaciones pasadas para comprender y responder a los usuarios.



Memorización: Los intercambios anteriores en una conversación entre un usuario y un chatbot se acumulan y se incorporan al prompt siguiente, permitiendo que el chatbot recuerde las entradas anteriores. Sin embargo, existe un límite en la longitud que puede alcanzar un prompt debido a la restricción de tokens en los modelos GPT-3, como ChatGPT, lo que puede limitar la capacidad de recordar conversaciones muy extensas.



Chatbot + base de conocimiento

Avances recientes en modelos de lenguaje grandes (LLMs) como GPT-3 y ChatGPT han generado interés en la industria tecnológica. Estos modelos son poderosos pero presentan desafíos como sesgos y alucinaciones. Uno de los usos prometedores de estos LLMs es el desarrollo de chatbots.



Chatbots basados en intención

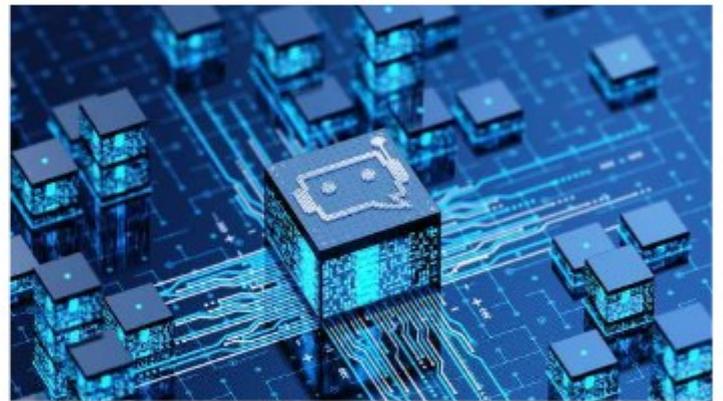
También llamados Chatbots tradicionales

Limitado a un set de respuestas.
Se utilizan en tareas fáciles como servicio al cliente.
Utiliza un enfoque basado en reglas.

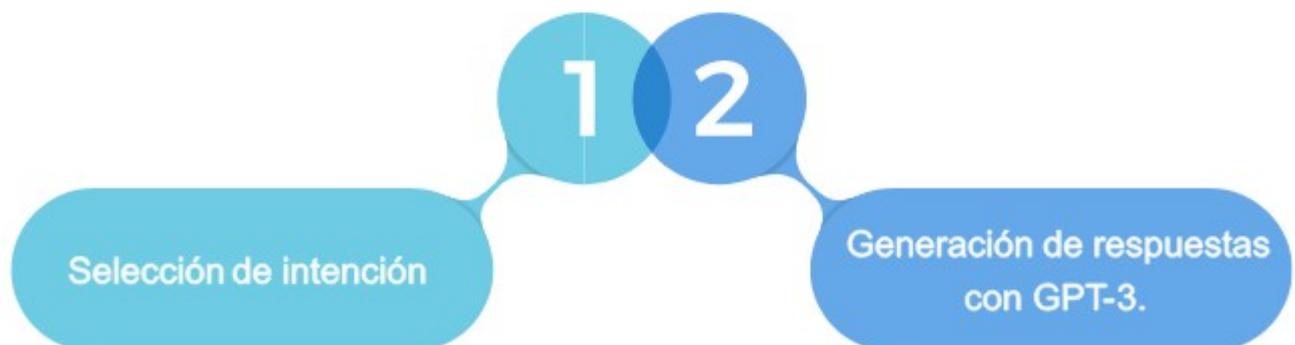
¿Cómo puede ayudar gpt-3?

GPT-3 puede ayudar reduciendo la cantidad de intenciones al asociar cada una con un documento en una Base de Conocimiento (KB). Esto permite respuestas más contextuales y menos intenciones específicas.

Los LLMs tienen límites de tamaño para prompts. Esto se debe a limitaciones computacionales. Futuros LLMs pueden no tener esta restricción.



¿Cómo podría funcionar un chatbot con gpt-3?



Generación de respuestas con gpt-3

Para crear la indicación (prompt), exploraremos las siguientes estrategias:

1- Role-prompting: una técnica para aumentar el conocimiento que asigna un papel específico a la inteligencia artificial.

Ejemplo: Como un chatbot avanzado, tu objetivo principal es ayudar a los usuarios de la mejor manera que puedas.

2-Incorporación de información relevante de la Base de Conocimiento (KB), es decir, el documento recuperado durante la búsqueda semántica.

Ejemplo: Imaginemos que, como resultado del proceso de búsqueda semántica, hemos recuperado un documento de nuestra Base de Conocimiento (KB). Todos estos documentos ofrecen información sobre el funcionamiento de "PetCareHub", un producto ficticio dedicada al cuidado de mascotas.



Iniciar sesión en PetCareHub desde el sitio web

Abre tu navegador web y ve al sitio web de PetCareHub.
Haz clic en el botón "Iniciar sesión" ubicado en la esquina superior derecha de la página.
En la página de inicio de sesión, introduce tu nombre de usuario y contraseña de PetCareHub.
Una vez que hayas ingresado tus credenciales, haz clic en el botón "Iniciar sesión".
Ahora deberías estar conectado a tu cuenta de PetCareHub.
Iniciar sesión en PetCareHub desde la aplicación móvil

Abre la aplicación de PetCareHub en tu dispositivo móvil.
En la página principal, toca el botón "Iniciar sesión" ubicado en la esquina inferior derecha.
En la página de inicio de sesión, introduce tu nombre de usuario y contraseña de PetCareHub.
Una vez que hayas ingresado tus credenciales, pulsa el botón "Iniciar sesión".
Ahora deberías estar conectado a tu cuenta de PetCareHub.

3-Incorporación de información relevante de la Base de Conocimiento (KB), es decir, el documento recuperado durante la búsqueda semántica.

Ejemplo: Podemos agregar su contenido dentro del prompt de esta manera.

Como un chatbot avanzado, tu objetivo principal es ayudar a los usuarios de la mejor manera que puedas

Iniciar sesión en PetCareHub desde el sitio web

Abre tu navegador web y ve al sitio web de PetCareHub.
Haz clic en el botón "Iniciar sesión" ubicado en la esquina superior derecha de la página.

En la página de inicio de sesión, introduce tu nombre de usuario y contraseña de PetCareHub.

Una vez que hayas ingresado tus credenciales, haz clic en el botón "Iniciar sesión".

Ahora deberías estar conectado a tu cuenta de PetCareHub.

Iniciar sesión en PetCareHub desde la aplicación móvil

Abre la aplicación de PetCareHub en tu dispositivo móvil.

En la página principal, toca el botón "Iniciar sesión" ubicado en la esquina inferior derecha.

En la página de inicio de sesión, introduce tu nombre de usuario y contraseña de PetCareHub.

Una vez que hayas ingresado tus credenciales, pulsa el botón "Iniciar sesión".

Ahora deberías estar conectado a tu cuenta de PetCareHub.

4- Inclusión de la pregunta formulada por el usuario como parte de la indicación.

Como un chatbot avanzado, tu objetivo principal es ayudar a los usuarios de la mejor manera que puedas

Iniciar sesión en PetCareHub desde el sitio web

Abre tu navegador web y ve al sitio web de PetCareHub.
Haz clic en el botón "Iniciar sesión" ubicado en la esquina superior derecha de la página.
En la página de inicio de sesión, introduce tu nombre de usuario y contraseña de PetCareHub.

Una vez que hayas ingresado tus credenciales, haz clic en el botón "Iniciar sesión".

Ahora deberías estar conectado a tu cuenta de PetCareHub.

Iniciar sesión en PetCareHub desde la aplicación móvil

Abre la aplicación de PetCareHub en tu dispositivo móvil.

En la página principal, toca el botón "Iniciar sesión" ubicado en la esquina inferior derecha.

En la página de inicio de sesión, introduce tu nombre de usuario y contraseña de PetCareHub.

Una vez que hayas ingresado tus credenciales, pulsa el botón "Iniciar sesión".

Ahora deberías estar conectado a tu cuenta de PetCareHub.

Chatbot: ¡Hola! ¿En qué puedo ayudarte?

USUARIO: No puedo encontrar el botón de inicio de sesión.

Chatbot:

La respuesta proviene del documento de base de conocimiento



Respuesta del Chatbot:

No hay problema! En el sitio web, el botón de inicio de sesión se encuentra en la esquina superior derecha de la página. En la aplicación móvil, el botón de inicio de sesión se encuentra en la esquina inferior derecha. Hágamelo saber si necesita más ayuda!

Problemas al generar respuestas con gpt-3

Si consultamos a GPT-3 la pregunta "¿La aplicación móvil es gratuita?" y le proporcionamos el documento de inicio de sesión como contexto, a menudo obtendremos una respuesta como "Sí, la aplicación móvil de PetCareHub es gratuita para descargar y usar", incluso si esta información no está presente en el contexto. La generación de información incorrecta es problemática para los chatbots de servicio al cliente.

GPT-3 rara vez produce información inexacta cuando la respuesta está disponible en el contexto.

Como un chatbot avanzado, tu objetivo principal es ayudar a los usuarios de la mejor manera que puedas

Iniciar sesión en PetCareHub desde el sitio web

Abre tu navegador web y ve al sitio web de PetCareHub.
Haz clic en el botón "Iniciar sesión" ubicado en la esquina superior derecha de la página.
En la página de inicio de sesión, introduce tu nombre de usuario y contraseña de PetCareHub.

Una vez que hayas ingresado tus credenciales, haz clic en el botón "Iniciar sesión".
Ahora deberías estar conectado a tu cuenta de PetCareHub.
Iniciar sesión en PetCareHub desde la aplicación móvil

Abre la aplicación de PetCareHub en tu dispositivo móvil.
En la página principal, toca el botón "Iniciar sesión" ubicado en la esquina inferior derecha.
En la página de inicio de sesión, introduce tu nombre de usuario y contraseña de PetCareHub.
Una vez que hayas ingresado tus credenciales, pulsa el botón "Iniciar sesión".
Ahora deberías estar conectado a tu cuenta de PetCareHub.

Chatbot: ¡Hola! ¿En qué puedo ayudarte?

USUARIO: ¿La aplicación móvil es gratuita?"

Chatbot: "Sí, la aplicación móvil de PetCareHub es gratuita para descargar y usar"

Referencias

- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language Models are Few-Shot Learners.
- Bonta, A. (2022). How to use OpenAI's ChatGPT to write the perfect cold email. <https://www.streak.com/post/how-to-use-ai-to-write-perfect-cold-emails>.
- Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., Barham, P., Chung, H. W., Sutton, C., Gehrmann, S., Schuh, P., Shi, K., Tsvyashchenko, S., Maynez, J., Rao, A., Barnes, P., Tay, Y., Shazeer, N., Prabhakaran, V., ... Fiedel, N. (2022). PaLM: Scaling Language Modeling with Pathways.
- Cobbe, K., Kosaraju, V., Bavarian, M., Chen, M., Jun, H., Kaiser, L., Plappert, M., Tworek, J., Hilton, J., Nakano, R., Hesse, C., & Schulman, J. (2021). Training Verifiers to Solve Math Word Problems.
- Jin, Y., Kadam, V., & Wanvarie, D. (2022). Plot Writing From Pre-Trained Language Models.
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Madotto, A., & Fung, P. (2022). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*. <https://doi.org/10.1145/3571730>
- Jurafsky, D., & Martin, J. H. (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. Prentice Hall.
- Karpas, E., Abend, O., Belinkov, Y., Lenz, B., Lieber, O., Ratner, N., Shoham, Y., Bata, H., Levine, Y., Leyton-Brown, K., Muhlgay, D., Rozen, N., Schwartz, E., Shachaf, G., Shalev-Shwartz, S., Shashua, A., & Tenenholz, M. (2022).
- Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2022). Large Language Models are Zero-Shot Reasoners.
- Lake, B. M., & Baroni, M. (2018). Generalization without Systematicity: On the Compositional Skills of Sequence-to-Sequence Recurrent Networks. <https://doi.org/10.48550/arXiv.1711.00350>
- Liu, J., Liu, A., Lu, X., Welleck, S., West, P., Bras, R. L., Choi, Y., & Hajishirzi, H. (2021). Generated Knowledge Prompting for Commonsense Reasoning.
- Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2022). Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. *ACM Computing Surveys*. <https://doi.org/10.1145/3560815>
- Nadeem, M., Bethke, A., & Reddy, S. (2021). StereoSet: Measuring stereotypical bias in pretrained language models. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 5356–5371. <https://doi.org/10.18653/v1/2021.acl-long.416>
- Nobel, P. S., & others. (2002). *Cacti: biology and uses*. Univ of California Press
- OpenAI. (2022). ChatGPT: Optimizing Language Models for Dialogue. <https://openai.com/blog/chatgpt/>. <https://openai.com/blog/chatgpt/>
- Shin, T., Razeghi, Y., Logan IV, R. L., Wallace, E., & Singh, S. (2020). AutoPrompt: Eliciting Knowledge from Language Models with Automatically Generated Prompts. *Proceedings of the 2020 Conference*