



Guía para crear “shots”

Se ha demostrado que es posible mejorar el desempeño de los **modelos de lenguaje (LLM)**, como **GPT-4**, potenciando su rendimiento en tareas generales utilizando unos cuantos ejemplos de lo que el usuario espera obtener como resultado del trabajo del LLM. Incluso, esta técnica puede ser comparada en resultados con las técnicas de ajuste fino (*fine-tuning*) más avanzadas (Brown et al., 2020):

Paso 1. Diseño

Diseña un *prompt* “base” para una tarea repetitiva que quieras automatizar o, por lo menos, dejar una parte de la carga al modelo de IA. Por ejemplo, para automatizar la generación de resúmenes con la extensión de un tuit, puedes usar este *prompt*:

“Eres un excelente sintetizador de información y experto en diversos temas. Te daré un texto en cualquier idioma y redactarás un resumen para mí. El resumen debe cumplir estas condiciones:

- Debe tener la extensión máxima de un tuit (140 caracteres)
- Mantén un lenguaje claro y simple
- Utiliza vocabulario y lenguaje tal que hasta una persona de 12 años pueda comprender.

##Texto a resumir:”

Paso 2. Iteración

Repite el paso 1 las veces que sean necesarias hasta que obtengas varios pares de *input (prompt)* - *output* (resultado). En este caso, serían pares texto-resumen.

Paso 3. Rediseño

Toma de nuevo tu *prompt* base y añade esos pares *input-output* exitosos para crear un nuevo *prompt*. Estos elementos se conocen como “shots”; es decir, ejemplos que se dan al LLM para que reconozca mejor el contexto de la solicitud. Se ha demostrado que a partir de 3 “shots”, los modelos ya no tienen mejoras considerables. Por lo que incluir de 2 a 3 “shots” es más que suficiente:

“Eres un excelente sintetizador de información y experto en diversos temas. Te daré un texto en cualquier idioma y redactarás un resumen para mí. El resumen debe cumplir estas condiciones:

- Debe tener la extensión máxima de un tuit (140 caracteres)
- Mantén un lenguaje claro y simple
- Utiliza vocabulario y lenguaje tal que hasta una persona de 12 años pueda comprender.

Aquí hay dos ejemplos de pares de inputs y outputs para que entiendas a lo que me refiero:

Ejemplo 1
Input 1: ...
Output 1: ...

Ejemplo 2
Input 2: ...
Output 2: ...

##Texto a resumir: {...}”

Paso 4. Almacenamiento

Por último, te recomendamos que guardes estos nuevos *prompts* en algún repositorio donde los puedas tener al alcance, y comiences a crear tu “**librería de prompts**”. Algunas opciones son:

- *Notion*
- *Google sites*
- *Text Blaze*
- Documentos .txt, .doc

Además, te recomendamos que también guardes en una hoja de cálculo esos partes *input/output* ya que podrán servirte si en el futuro quieres hacer un chatbot con *fine-tuning*.

Referencias:

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Amodei, D., et al. (2020). *Language models are few-shot learners*. Recuperado de <https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf>