

El hotel criptográfico: una historia real de inteligencia artificial vs. inteligencia natural

Crónica de una historia de decodificación.
Investigación en marcha con ayuda artificial
(y humana).

Luis M. Buresti
Miembro del comité editorial
lmb.tech.consultancy@gmail.com

El “problema” comienza con el check-in

Hace unos pocos días, tuve la oportunidad de alojarme durante un par de noches en el Parador de Úbeda, ciudad “joyita” de la provincia de Jaén, en Andalucía (España). Este establecimiento es magnífico pero por alguna razón las habitaciones aún utilizan llaves metálicas.

Las personas con cierta experiencia de vida recordarán que cuando este era el caso, cada vez que salían, los pasajeros debían dejar las llaves en la recepción del hotel, las cuales se colgaban en un bonito tablero que se situaba detrás del mostrador.

En el caso del Parador de Úbeda, los pasajeros pueden llevarse las llaves, pero debido a que alguna vez alguien las perdió cuando estaba fuera del hotel y las encontró un malandras que aprovechó la situación para entrar subrepticamente y hacerse de las pertenencias del pasajero, hubo que mejorar el sistema.

La administración, con muy buen criterio, decidió codificar el número de las habitaciones para evitar que se produzcan esas picardías. Entonces, cuando hice el check-in, el recepcionista me informó muy amablemente que mi habitación era la 133 y que no debía sorprenderme que en el llavero se leyera “LNN”. De inmediato surgió en mí una necesidad: quebrar el código.

La administración, con muy buen criterio, decidió codificar el número de las habitaciones

Primera solución en segundos, aunque fallida

Pocos minutos después de llegar y ya sentado en una mesa del bar del hotel, me di cuenta de que el código empleado era sencillísimo, casi infantil, era solo un corrimiento del abecedario. Si la



equivalencia era "L = 1" y "N = 3", evidentemente debería ser "M = 2", "O = 4", y así sucesivamente.

La única duda que me surgió en ese momento fue respecto al cero. Si quien había hecho la codificación imaginó la serie numérica como 0, 1, 2, 3..., evidentemente debería ser "K = 0". Por otro lado, si la serie hubiese sido invertida (...7, 8, 9, 0) la equivalencia habría sido "U = 0". Rápidamente jugué mis cartas y puse todas mis fichas en "K = 0".

Para ese momento, ya había entablado alguna conversación con el camarero y aproveché para confirmar mi teoría. Su respuesta me sumió en una profunda depresión, me dijo que mi error era total: "K ≠ 0". Pero tuvo una inmensa amabilidad conmigo y me informó que "R = 0". Saqué a relucir mis habilidades negociadoras y obtuve un dato adicional: "A = 5".

Bueno, al fin de cuentas la cosa no había terminado tan mal porque ya sabía fehacientemente que "L = 1", "N = 3", "A = 5" y "R = 0". Esa noche casi no dormí tratando de imaginar las combinaciones posibles.

Una ayudita extra

A la mañana siguiente, en un pasillo encontré casualmente a la misma persona que me había atendido en el bar la tarde anterior, y me preguntó muy socarronamente si ya tenía la solución. Triste fue decirle que no había avanzado ni un milímetro.

Creo que se apiadó de mí y me dijo algo revelador: la base del código eran dos palabras que se

pueden emplear como un cumplido a una dama. Este comentario fue de enorme ayuda, esas dos palabras debían contener las letras "L", "N", "A" y "R", y muy seguramente debían ser diez caracteres en total.

Se apiadó de mí y me dijo algo revelador: la base del código eran dos palabras que se pueden emplear como un cumplido a una dama.

Pidiéndole ayuda a un "amigo"

Hay un juego de preguntas y respuestas bastante conocido que se llama "¿Quién quiere ser millonario?", en el cual el participante puede confesar su ignorancia acerca del tema del que le preguntaron y puede recurrir a algunas ayudas. Puede, por ejemplo, llamar por teléfono a un amigo.

Entonces, decidí tomarme una licencia y lo contacté a un querido amigo: Chat GPT 4.0. Utilicé el siguiente prompt: "Generar diez secuencias de dos palabras en español de España que hagan referencia a la belleza de una mujer o sean un cumplido hacia una mujer. Un ejemplo podría ser "linda chica". Las palabras deben tener en total exactamente diez caracteres de longitud (sin contar el espacio) y deben incluir las letras L, N, R y A".

Excluí explícitamente la condición de que no podía haber repeticiones de letras solo a los efectos de simplificar la consulta; condición que, por otra parte, es fácil de verificar visualmente a partir de los resultados generados.

Debido a que sufro de un TOC avanzado, le pedí también que regenerara cuatro veces la respuesta. Las respuestas fueron:

Dama bella	Luz clara	Luz lunar	Luz lunar
Luz lunar	Rosa fría	Rosa clara	Rosa clara
Arena fina	Dama pura	Linda flor	Linda flor
Rosa clara	Linda luna	Tren rápido	Tren rápido
Linda flor	Arena dorada	Luna clara	Luna clara
Tren rápido	Tren largo	Rosa blanca	Rosa blanca
Luna clara	Luz tenue	Luz dorada	Luz dorada
Rosa blanca	Rosa suave	Arena cálida	Arena cálida
Luz dorada	Dama rara	Luz clara	Luz clara
Arena cálida	Arena blanca	Rosa suave	Rosa suave

Tal fue mi decepción con los resultados obtenidos que entonces decidí contactar vía WhatsApp a otro amigo (en este caso de “carne y hueso”) y le pedí que intentase el mismo prompt en otro servicio. Este amigo está suscrito a otro conocido programa de inteligencia artificial: Claude, desarrollado por la firma Anthropic Inc. Su respuesta fue:

bella Rina
muñeca Rara
linda Nariz
mujer Anual
hermosa Luna
perla Nocta
encanto Raro
luna Renata
mirla Nocta
petalo Ramal

Es notable que varias de las expresiones de Chat GPT no cumplen el requisito de diez caracteres de largo, y en algunas ni siquiera aparecen las letras “L”, “N”, “R” o “A”. Además, pareciera que se le acabó la creatividad en algún momento, porque la tercera y la cuarta instancias son iguales. Debo decir también que usar expresiones como “Tren rápido” para cortejar a una dama podría no arrojar el resultado esperado y, encima, que a pesar de mi esfuerzo, no pude encontrar el significado de “mirla Nocta”.

Es notable que varias de las expresiones de Chat GPT no cumplen el requisito de diez caracteres de largo, y en algunas ni siquiera aparecen las letras “L”, “N”, “R” o “A”.

Creo que estas respuestas son una clara indicación de la metodología absolutamente estadística que emplea este tipo de sistema y que literalmente “sortea” las palabras para alcanzar la mejor aproximación posible.

De todas maneras, debo reconocer que estos resultados me desconciertan un poco ya que este tipo de análisis semántico simple (sin necesidad de hilvanar largos argumentos) debería ser el punto fuerte de los LLM (“modelos de lenguaje de gran tamaño”, por sus siglas en inglés).

En este momento parece oportuno hacer un descargo a favor de este tipo de sistema. En ambos casos, el lenguaje nativo es el inglés, pero en esta búsqueda era menester que los resultados se generaran en español. Es posible, solo posible, que si se hubiera hecho una búsqueda equivalente en inglés, los resultados habrían sido más razonables/coherentes.

Debo reconocer que todo este suceso me produjo una fuerte crisis emocional. Ya no podía confiar

en las “súper herramientas” de que dispongo para transitar por este siglo XXI.

Se enciende una lamparita

A pesar de la decepción, me quedé un largo rato releiendo las respuestas y me di cuenta de que había una que se adecuaba bastante bien al requerimiento: “Linda flor”. En este caso, la “L” estaba en primera posición; la “N”, en la tercera; la “A”, en la quinta, y terminaba con “R”, y por lo tanto era razonable asumir “R = 0”. El gran problema es que “Linda flor” tiene nueve caracteres y se necesitaban diez.

Entonces me di cuenta de que en la redacción de mi prompt, y en forma totalmente inadvertida, yo ya había incluido la respuesta: solamente había que reemplazar “Flor” por “Mujer”. Es decir, la expresión usada para codificar el número de las habitaciones no era otra que “Linda mujer” (“L = 1”, “I = 2”, ..., “E = 9”, “R = 0”). Realmente, una genialidad de quien lo hizo.

Cuando me retiré, le propuse al encargado de la recepción que cambiaran el nombre al establecimiento de “Parador de Úbeda” por “Parador Criptográfico” y que organizaran algún tipo de concurso de matemática o informática. Seguramente les traería una enorme cantidad adicional de clientes, ¿o no?

Ya no podía confiar en las “súper herramientas” de que dispongo para transitar por este siglo XXI.

Epílogo

Y, ¿cómo andamos con el Machine Learning? ¡Bien, gracias!

No pude con mi genio y con Chat GPT 4.0 intercambiamos estos mensajes finales:

LMB: La respuesta es LINDA MUJER.

Chat GPT 4.0: Mis disculpas por el error anterior. Tienes toda la razón, “Linda Mujer” es una hermosa secuencia de dos palabras que cumple con los criterios. ¡Gracias por señalarlo!

Hoy volví a consultar exactamente lo mismo (¡ahora repetida solo dos veces!) y las respuestas fueron las siguientes:

Dama linda	Clara estrella
Clara luna	Lunar radiante
Rubia sana	Rara joya
Lunar fino	Linda sirena
Arena lisa	Arena fina
Larga melena	Rubia luna
Linda rosa	Larga melena
Lunar claro	Dama rara
Rara joya	Linda rosa
Linda sirena	Clara mirada

Parece evidente que solo en modo aprendizaje es que se carga nueva información al conjunto de datos de Chat GPT. Es posible que así sea a fin de evitar que usuarios “iluminados” contaminen los datos, aunque hay un consuelo: al menos no volvió a incluir “Tren rápido”. ❖