

Recogida de la dieta con chatGPT

Olivier Nuñez (onunez@isciii.es)

Seminario del CNE, 1 de Junio de 2023

Motivación

Proyecto ALEVINT (Acción Estratégica en Salud, IP: Emma Ruiz)

Descomposición nutricional estandarizada de cuestionarios dietéticos

Entre otros objetivos:

- Recogida de información dietética a gran escala
- Mejorar participación y adherencia al cuestionario
- Procesamiento estandarizado de la información
- Bajo coste

Estructura de la presentación

1. ¿Porqué y cómo recoger la dieta?
2. El ChatBot como entrevistador dietético
3. Implementación

**¿Porqué y cómo recoger la
dieta?**

Importancia de la dieta en la salud

“Que tu medicina sea tu alimento!”, Hipócrates

- **Primeras evidencias:** enfermedades relacionadas con la carencia de nutrientes (escorbuto, beriberi, pelagra).
- **Enfermedades infecciosas.** Una buena dieta contribuye a una mejor respuesta inmunitaria frente a las infecciones.
- **Enfermedades crónicas.** Papel de la dieta respaldado por varios estudios:
 - **E. Cardiovasculares:** PREDIMED, Framingham, ...
 - **Cancer:** Nurse Health Study, EPIC, ...

¿Porqué recoger la ingesta de alimentos?

- Analizar patrón de ingesta de alimentos, energía y nutrientes en la población
- Comprender el papel de los nutrientes en la prevención de la salud
- Evaluar la eficacia de intervenciones dietéticas

¿Cómo recoger la ingesta de alimentos?

Metodos estandares

- Diario dietético (prospectivo abierto) : el *gold standard*
- Cuestionarios de frecuencia de consumo de alimentos (retrospectivo cerrado)
- Recuerdo de 24h (retrospectivo abierto)

Recuerdo de 24h

Apto para una conversación

MEDIA MAÑANA	Hora:	Lugar:
Menú y Proceso culinario		Alimentos (calidad y cantidad)
COMIDA	Hora:	Lugar:
Menú y Proceso culinario		Alimentos (calidad y cantidad)
Macarrones con chorizo gratinados		Macarrones (1 ración mediana, peso en crudo: 60g) Cebolla (1/3 de cebolla mediana) Chorizo (3 rodajas) Tomate frito (2 cucharadas soperas)
Pollo asado con ensalada de lechuga y tomate		Queso emmenthal rallado (ración) Pollo (ración) Lechuga iceberg (ración)
Melocotón en almíbar		Tomate (unidad mediana)
Agua y pan		Melocotón en almíbar (ración)
		Bebidas: agua (2 vasos) Pan: Pan blanco de barra (2 rebanadas) Aceite (tipo): Aceite de oliva virgen para ensalada (cucharada sobera rasa)

Recuerdo de 24h

Limitaciones: las de información autoreportada

- **Sesgo de memoria:** subestimación de la ingesta especialmente en ancianos y niños
- **Variabilidad:** la ingesta de alimentos puede variar sustancialmente de un día a otro
- **Influencia social:** infranotificación de alimentos no saludables
- **Imprecisión** en la evaluación de las porciones y descripción de la preparación de los alimentos

Recuerdo de 24h

Ventajas

- **Validez:** apto para estimar la ingesta de energía y nutrientes
 - Incluido en el NHANES (National Health and Nutrition Examination Study)
 - Siendo retrospectivo no altera el consumo de alimentos
- **Precisión:** constituye una buena estimación de la ingesta habitual mediante recuerdos en serie (2 o 3)
- **Viabilidad:** fácil de implementar, rápido y con un alto nivel de aceptación

ChatBot

Primer chatBot de la historia

ELIZA (Joseph Weizenbaum, 1966)

```
Welcome to

EEEEEE LL      IIII  ZZZZZZ  AAAAA
EE      LL      II    ZZ     AA   AA
EEEEEE LL      II    ZZZ    AAAAAA
EE      LL      II    ZZ     AA   AA
EEEEEE LLLLLL IIII ZZZZZZ  AA   AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:
```

Procesamiento del lenguaje natural

El humano ya no tiene el monopolio de su lenguaje!

ChatGPT es un *chatBot* basado en un modelo de procesamiento del lenguaje natural (GPT, 2018):

1. Análisis formal del lenguaje (correlación entre tokens)
2. Entrenamiento supervisado (por la propia IA) sobre grandes muestras de texto (se ha leído todo Internet!)
3. Predicción de oraciones moderada por una capa de control (evitar barbaridades!).

Limitaciones

- **Sesgos:** basado en una información en parte errónea
- **Inconsistencia:** sensibilidad a la formulación de la consulta
- **Falta de prudencia:** incluso en caso de ambigüedad, genera siempre una respuesta que podría ser incorrecta o inadecuada (un chatbot no sabe si miente!).
- (Falta de conocimiento actualizado)

chatGPT como entrevistador dietetico

Lo que importa es que sea buen conversador!

- Supera el test de Turing (más de 12 fotogramas por segundos!)
- Evita el sesgo del entrevistador (no tiene juicio moral!)
- Se adapta en tiempo real y se le puede adiestrar
- Mejora potencialmente la adherencia al cuestionario
- Económico: 0,20 céntimos por cada 100.000 tokens

Validación


TODO

- **Viabilidad:** sencillo y aceptado por usuarios (✓)
- **Validez:** la del R24h
- **Fiabilidad** (ToDo)
 - *Estabilidad:* entre medidas repetidas en el tiempo
 - *Concordancia:* nivel de acuerdo entre evaluadores distintos

Implementación

Contexto

Playground

Load a preset... 

Save

View code

Share



SYSTEM


Eres un asistente que realiza una encuesta de recuerdo de 24 horas mediante una conversación informal con un usuario. Tus preguntas trataran de determinar con la mayor precisión posible el tipo y la cantidad de los alimentos consumidos durante el día anterior.

USER



Soy Olivier

ASSISTANT

¡Hola Olivier! Soy un asistente virtual que tiene como objetivo ayudarte a recordar los alimentos que has consumido durante el día anterior. ¿Te gustaría responder algunas preguntas al respecto?

 Add message

Mode

 Chat Beta 

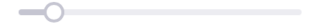
Model

gpt-3.5-turbo 

Temperature 1



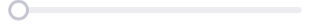
Maximum length 256



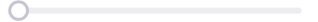
Top P 1



Frequency penalty 0



Presence penalty 0



Adiestramiento

- El modelo general no se adapta bien al papel de entrevistador dietético:
 - Varias preguntas simultaneas y muy largas (ise enrolla!)
 - Falta de precisión en las preguntas
- Dos métodos de adiestramiento:
 - Contexto más preciso (ino hace mucho caso!)
 - Ejemplos (fine-tuning):

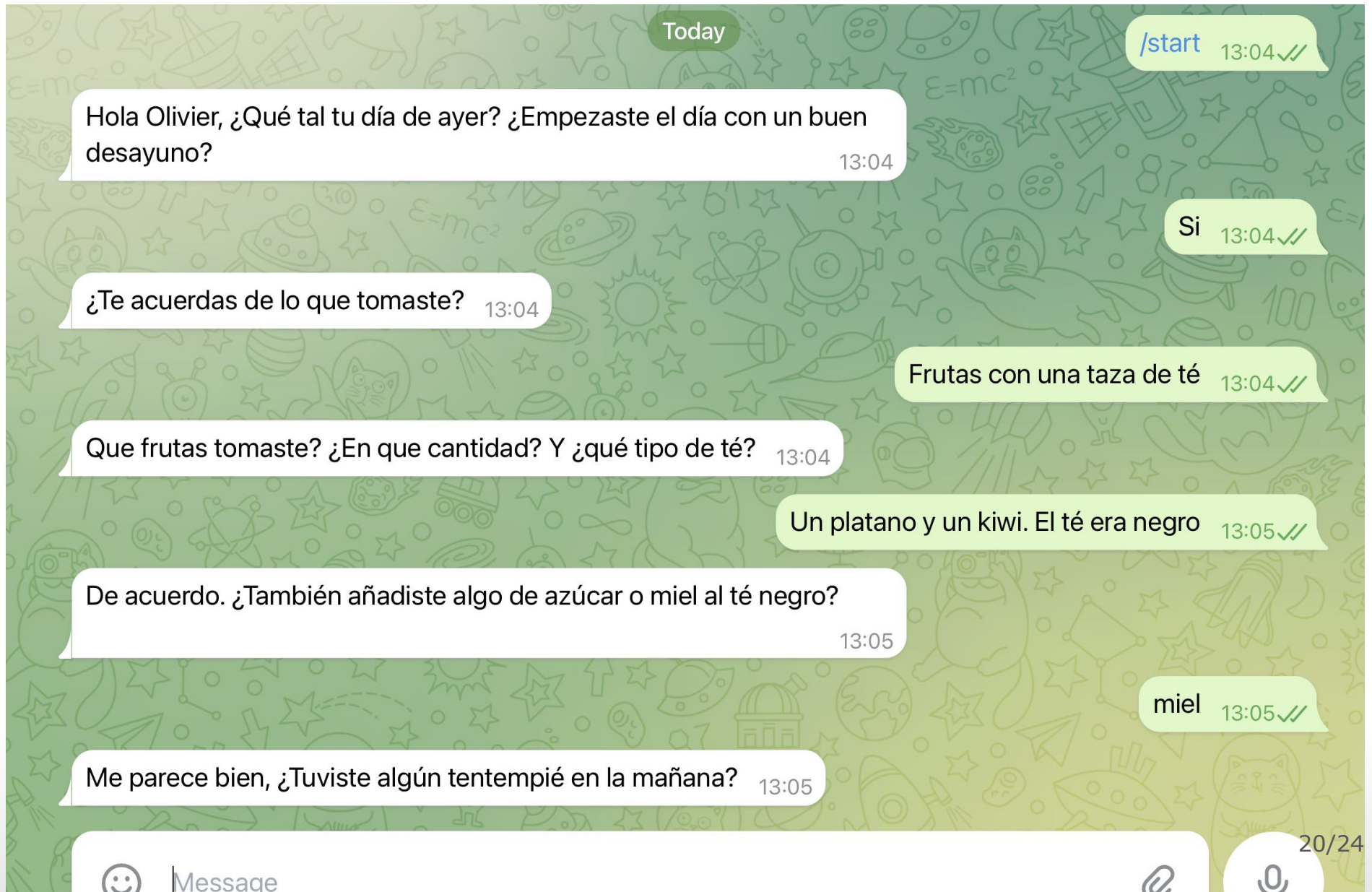


chat_07.txt

```
user:Soy María
assistant:Hola María| ¿Qué tal tu día de ayer? ¿Empezaste el día con un buen desayuno?
user:Si
assistant:¿Te acuerdas de lo que tomaste?
user:Un donuts de chocolate y a media mañana un pincho de tortilla
assistant:¿El donut era entero de chocolate?
user:No, era mitad chocolate
assistant:¿Qué tamaño tenía el pincho?
user:Era menos de un cuarto de la tortilla
assistant:¿lo has acompañado con pan?
user:Si
assistant:¿pan de baguette? ¿Cuánta cantidad de pan mas o menos?
user:Pan de barra, dos rebanadas
assistant:¿Lo acompañaste con algo de beber?
```

Llevando la entrevista al móvil

ChatGPT en Telegram



Procesamiento de la conversación

SYSTEM

Una hoja de cálculo de 5 columnas con los alimentos consumidos por el usuario:

user:Soy María

assistant:Hola María ¿Qué tal tu día de ayer?

¿Empezaste el día con un buen desayuno?

user:Si

assistant:¿Te acuerdas de lo que tomaste?

user:Un donuts de chocolate y a media mañana un pincho de tortilla

assistant:¿El donut era entero de chocolate?

user:No, era mitad chocolate

assistant:¿Qué tamaño tenía el pincho?

user:Era menos de un cuarto de la tortilla

assistant:¿lo has acompañado con pan?

user:Si

assistant:¿pan de baguette? ¿Cuánta cantidad de pan mas o menos?

user:Pan de barra, dos rebanadas

assistant:¿Lo acompañaste con algo de beber?

user:La tortilla con una Coca cola zero

assistant:¿Una lata, un vaso o cuánto?

user:Una botella de bar

assistant:¿Pequeña o grande?

user:Grande

assistant:¿Grande de dos litros?

user:No, de 500ml

Alimento|Tipo|Unidad|Cantidad|Comida

ASSISTANT

Donuts de chocolate|Dulce|Unidad|1/2|Desayuno
Pincho de tortilla|Salado|Unidad|Menos de 1/4|Media mañana
Pan de barra|Carbohidrato|Rebanada|2|Media mañana
Coca cola zero|Bebida|Botella|500ml|Media mañana

⊕ Add message

Implementación en R

Dialogando con interfaces

- **Dialogo con la Interfaz de openAI:** paquete [openai](#) de R
- **Dialogo con la Interfaz de Telegram:** paquete [telegram.bot](#) de R
- **Sencillez:** en total apenas 100 lineas de código en R !
- **Requisitos:** Ambas Interfaces requieren claves (API key) !

Conclusiones

- **Limitaciones:**

- Imprecisión en las preguntas (en vía de subsanación con GPT4)
- Sólo implementado en Telegram
- Confidencialidad:
 - chatGPT: no sabe con quien habla.
 - Telegram: mensajes cifrados (en principio de alta seguridad)

- **Fortalezas:**

- Facilita el trabajo de campo: de fácil difusión y barato
- Gran potencialidad: la IA mejora cada día
- Aumento de la participación: chatear es de moda

- **Perspectivas:**

- Descomposición (vía ALEVINT) en tiempo real del recuerdo en nutrientes y energía.
- Posibilidad de insertar descripciones visuales o auditivas de los platos en el chat

Referencias

- *Nutritional Epidemiology*, Willett WC. Oxford University Press, 2012.
- *The dietary assessment of individuals: methods, accuracy, new techniques and recommendations*, Bingham SA. *Nutr Abs Rev*;57:705-742, 1987.
- *OpenAI API References*: <https://platform.openai.com/docs/api-reference>
- *Telegram Bot API References*: <https://core.telegram.org/api>