

\_4bucles/\_12memoizaciointexto.py

```
1 # Ejercicio 12: Ejercicio de Memoización en Análisis de Texto
2
3 # Imagina que estás trabajando en un sistema de análisis de texto que requiere calcular la
  frecuencia de ocurrencia de palabras en un conjunto de documentos. Implementa una función
  llamada calcular_frecuencia_palabras que tome como entrada un texto y devuelva un diccionario
  que muestre la frecuencia de cada palabra en el texto.
4
5 # La función debe ser capaz de manejar textos y ser insensible a mayúsculas/minúsculas (por
  ejemplo, "Hola" y "hola" se consideran la misma palabra).
6 # Se deben excluir las palabras comunes (artículos, preposiciones, etc.) que no aportan
  información relevante al análisis.
7 # Utiliza memoización para evitar recalcular la frecuencia de palabras para el mismo texto.
8
9 import functools
10 import re
11 import time
12 @functools.lru_cache(maxsize=None)
13 def calcular_frecuencia_palabras(texto):
14     #Preprocesamiento del texto
15     texto=texto.lower()
16     texto=re.sub(r'^a-z\s|', '', texto) #Eliminar puntuación o caracteres no deseados
17     palabras=texto.split()
18
19     #Excluir palabras comunes
20     palabras_comunes = set(["el", "la", "los", "las", "de", "en", "y", "a", "con", "es", "un",
  "una", "para"])
21
22     palabras=[palabra for palabra in palabras if palabra not in palabras_comunes]
23
24     #Conteo de frecuencia
25     frecuencia={}
26     for palabra in palabras:
27         frecuencia[palabra]=frecuencia.get(palabra,0)+1
28     return frecuencia
29
30 # Ejemplo de uso
31 texto_1 = """
32 En un lugar de la Mancha, de cuyo nombre no quiero acordarme,
33 no ha mucho tiempo que vivía un hidalgo de los de lanza en astillero,
34 adarga antigua, rocín flaco y galgo corredor.
35 Una olla de algo más vaca que carnero, salpicón las más noches,
36 duelos y quebrantos los sábados, lentejas los viernes,
37 algún palomino de añadidura los domingos, consumían las tres partes de su hacienda.
38 El resto della concluían sayo de velarte, calzas de velludo para las fiestas con sus
  pantuflos de lo mismo,
39 los días de entre semana se honraba con su vellorí de lo más fino.
40 Tenía en su casa una ama que pasaba de los cuarenta y una sobrina que no llegaba a los
  veinte,
41 y un mozo de campo y plaza que así ensillaba el rocín como tomaba la podadera.
42 Frisaba la edad de nuestro hidalgo con los cincuenta años; era de complexión recia, seco de
  carnes,
43 enjuto de rostro, gran madrugador y amigo de la caza.
```

```
44 Quieren decir que tenía el sobrenombre de Quijada o Quesada (que en esto hay alguna
    diferencia en los autores que deste caso escriben),
45 aunque por conjeturas verosímiles se deja entender que se llama Quijana;
46 pero esto importa poco a nuestro cuento: basta que en la narración dél no se salga un punto
    de la verdad.
47 """
48 texto_2 = "Otro ejemplo de Análisis de texto con palabras repetidas. Ejemplo y texto se
    repiten."
49
50 inicio_1=time.time()
51 frecuencia_sin_memo_1=calcular_frecuencia_palabras(texto_1)
52 frecuencia_sin_memo_2=calcular_frecuencia_palabras(texto_2)
53 final_1=time.time()
54 inicio_2=time.time()
55 frecuencia_con_memo_1=calcular_frecuencia_palabras(texto_1)
56 frecuencia_con_memo_2=calcular_frecuencia_palabras(texto_2)
57 final_2=time.time()
58
59 print("Frecuencia de palabras para texto_1 sin memoizacion: ",frecuencia_sin_memo_1)
60 print("Frecuencia de palabras para texto_2 sin memoizacion: ",frecuencia_sin_memo_2)
61 print("Frecuencia de palabras para texto_1 con memoizacion: ",frecuencia_con_memo_1)
62 print("Frecuencia de palabras para texto_2 con memoizacion: ",frecuencia_con_memo_2)
63
64 print("Tiempo sin memoizacion: ",(final_1-inicio_1))
65 print("Tiempo con memoizacion: ",(final_2-inicio_2))
```