

Problema C: Analizando Tweets



ProgramaMe Regional Online Valencia 2018 - 2019 - CEEDCV (Valencia)



Desde ProgramaMe somos muy curiosos y nos gusta saber si la gente está hablando de nosotros en Twitter (esperemos que bien ;). Para ello estamos analizando mogollón de tweets a mano, pero es una tarea agotadora y entre eso y que somos un poco vagos, hemos pensado en pedirnos vuestra inestimable ayuda.

Para analizar los tweets hemos realizado el siguiente proceso:

- 1) En primer lugar leemos un *tweet base* que sabemos a ciencia cierta qué está hablando de ProgramaMe. Del tweet base, asignamos a cada palabra (sin distinguir mayúsculas de minúsculas) un número según el número de veces que aparece en los tweets a analizar. Dicho número será la puntuación asociada a esa palabra.
 - a) Por ejemplo, si una palabra aparece una vez, su puntuación será 1, si aparece dos veces será 2, si aparece tres veces será 3, etc...
- 2) Tras ello, leeremos un número. Este número será un umbral que deberá ser alcanzado o superado por la puntuación de un tweet para que consideremos que dicho tweet habla de ProgramaMe.
- 3) A continuación, analizaremos una serie de *tweets desconocidos* que no sabemos si hablan de ProgramaMe o no. A cada tweet le daremos una puntuación de la siguiente forma:
 - a) Inicialmente la puntuación del *tweet desconocido* es cero.
 - b) Analizaremos cada palabra del *tweet desconocido*. Si alguna palabra coincide con alguna de las palabras del *tweet base* (sin distinguir mayúsculas y minúsculas), le sumamos la puntuación de dicha palabra. Si en el tweet

desconocido alguna palabra del tweet base se repite más de una vez, se volverá a sumar su puntuación.

- 4) Sabiendo la puntuación obtenida, compararemos contra el umbral que se nos ha proporcionado y si lo iguala o supera diremos que dicho tweet habla de ProgramaMe. En caso contrario, diremos que no habla de ProgramaMe.

Por ejemplo, si tenemos como *tweet base*.

ProgramaMe me encanta y me fascina

La palabra *ProgramaMe* tiene una puntuación de 1, la palabra *me* de 2 (aparece 2 veces), la palabra *y* de 1, la palabra *encanta* de 1 y la palabra *fascina* de 1.

Supongamos que tenemos un umbral de 3.

Y si un *tweet desconocido* dice:

Programame es emocionante y me encanta

Aparece 1 vez la palabra *Me* y 1 vez la palabra *ProgramaMe*, por lo cual dicho tweet alcanza una puntuación de 3. Al ser mayor o igual que el umbral establecido, consideraremos que se está hablando de ProgramaMe.

Entrada

En primer lugar, una frase indicando el tweet base. Las palabras estarán separadas por un espacio y sólo incluirán mayúsculas, minúsculas y números.

Tras ello una línea con un número U indicando el umbral exigido para saber si un tweet habla de ProgramaMe o no.

- $1 \leq U \leq 100000$

Tras ello, una línea con un número N indicando cuántos tweets debe analizar.

- $1 \leq N \leq 1000$

Durante las siguientes N líneas, cada línea contendrá un tweet. Cada tweet únicamente estará formado por mayúsculas, minúsculas, números y espacios. Las palabras se separarán por un único espacio.

Aunque en Twitter los tweets son de 280 caracteres máximo, para nuestro problema utilizaremos “tweets especiales” donde cada tweet poseerá un máximo de 2000 caracteres.

Salida

Se mostrarán N líneas, una por cada tweet desconocido.

La línea indicará *SI* en el caso de que el tweet esté hablando de ProgramaMe y *NO* en caso contrario.

Ejemplo de entrada

ProgramaMe me encanta y me fascina

3

8

Programame es emocionante y me encanta

Programa me

Programame me programa el alma

Programando voy programando vengo

PROGRAMAME ME ENCANTA Y ME FASCINA

Me Me

ProgramaMe es lo mejor desde Programame

ProgramaMe

Ejemplo de salida

SI

NO

SI

NO

SI

SI

NO

NO