

Clasificación y Representación de Documentos - Tweets de Twitter:

- 1.- Leer los datos de los tweets en tweets.txt.
- 2.- Crear un corpus C de documentos utilizando estos datos.
- 3.- Preprocesar los datos para la limpieza necesaria para obtener información sobre los datos.
- 4.- Representar los documentos en forma de matriz.
- 5.- Encontrar las palabras más relevantes de todos los documentos.
- 6.- Visualizar las frecuencias de las palabras encontradas en la matriz de términos de documentos usando un wordcloud.
- 7.- ¿Qué podemos deducir sobre los tweets al observar el wordcloud?
- 8.- Encontrar los tweets más similares de los datos de Twitter.