

2.2 Titulares de Fake News

Los *Fake News* son cada vez más populares, si bien muchos de ellos son escritos manualmente para difundir contenido falso, otros son elaborados de manera automatizada. Entender cómo se generan, permite identificarlos de forma más sencilla. Por lo cual, realizará cada una de las distintas etapas que se ven involucradas en dicha labor. De esta manera, se requiere que descargue una cantidad n de titulares de un portal de noticias como [BBC Mundo](#), mediante técnicas de *web-scraping*. Esta información debe ser almacenada en un archivo .csv autogenerado por el programa, que contenga la fecha de la noticia y el titular, y cualquier otro dato, de ser necesario. Así mismo, con los datos recopilados debe generar su propio titular falso utilizando Cadenas de Markov. Finalmente, retorne gráficos que permitan realizar un análisis de la información recopilada:

- Distribución de frecuencia de las palabras utilizadas en todos los enunciados. Tener en cuenta una etapa previa de limpieza de la información para evitar las stop-words.
- Palabra más usada por fecha de publicación de la noticia. Puede escoger el tipo de gráfico que considere más conveniente.
- Cantidad de artículos publicados por fecha.
- Nube de palabras más utilizadas.
- Distribución de los 10 bigramas y trigramas más frecuentes.
- Distribución de los tipos de palabras más frecuentes en los enunciados (adjetivos, verbos, pronombres, adverbios, etc.).

Genere 30 titulares aleatorios, basado en la información recopilada, y observe si los resultados se mantienen al realizar nuevamente el análisis anterior. Los resultados y las conclusiones obtenidas deben ser incluidas en el informe.

Nota: La solución debe implementarse utilizando las siguientes librerías de Python: *Pandas*, *seaborn*, *matplotlib*, *bs4*, *urllib*, *nlTK*. Se recomienda utilizar técnicas de procesamiento de lenguaje natural (NLP, por sus siglas en inglés) para llevar a cabo los análisis de palabras.