

# Práctica 1A

Guillermo Ruiz

Crea un notebook para preprocesar el texto y convertirlo a minúsculas, quita acentos y los siguientes caracteres:

" ; : , . \ \ - \ " ' / ( ) [ ] ¿ ? ¡ ! { } ~ < > | « » -- ' "

Reemplaza las tabulaciones y saltos de línea con un espacio.

2. Calcula las frecuencias a nivel de un carácter, de los documentos text\_1.txt al text\_5.txt y genera los histogramas para cada archivo. Para generar las gráficas, ordena los caracteres de acuerdo a su frecuencia, de mayor a menor frecuencia en el histograma.

Los archivos son:

- text\_1.txt
- text\_2.txt
- text\_3.txt
- text\_4.txt
- text\_5.txt

3. Interpreta cada histograma en referencia a qué caracteres destacan y dar indicios de por qué ocurren esas observaciones.