

# Práctica 1C: Cálculo de entropía

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.

1. Calcula la entropía global a nivel de carácter de los documentos text\_1 al text\_5 de manera independiente. No tomar en cuenta el espacio para el cálculo de la entropía.
2. Calcula la entropía global a nivel de palabra de los documentos libro\_1 y libro\_2 de manera independiente, primero sin quitar stopwords y luego quitándolas.

Archivos:

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