

Taller 1
ICI612 Electivo Profesional
Text Mining

Eliana Providel Godoy – eliana.providel@uv.cl

12 de septiembre, 2025

Información general:

- Taller para ser realizado en grupo de dos personas.
- Utilizar Jupyter notebook y dataset de presentación 1.
- Fecha de Entrega: miércoles 24 de septiembre, hasta las 23:59 hrs. Por medio del Aula Virtual.
- Enviar respuestas en documento word o pdf, y código en jupyter (entregar jupyter con las salidas, es decir no limpiar celdas)

Enunciado:

- 1) (15%) Implemente la ley de Zipf utilizando el texto de su dataset. Muestre el gráfico obtenido.
- 2) (20%) Dado el tema de su dataset, ¿Qué puede concluir a raíz del gráfico? Obs. Recuerde la definición de Text Mining vista en clases para dar respuesta.
- 3) (30%) Implemente tres características del texto de las expuestas en la Presentación 1. Puede utilizar librerías como NLTK o similares.
- 4) (35%) Para cada etiqueta del dataset y para cada característica implementada ¿Qué puede concluir? ¿Qué patrones caracterizan cada etiqueta del dataset? Presente al menos 5 ejemplos por cada característica/etiqueta.

Rúbrica de evaluación

Ítem	¿Qué se evaluará?	No logrado	Medianamente logrado	Logrado
	Resultado y/o completitud de la respuesta.	La respuesta tiene errores o se encuentra muy incompleta, sin embargo presenta de forma muy general la respuesta.	La respuesta presenta errores parciales, o se encuentra incompleta de forma parcial, sin embargo presenta detalles y aspectos correctos de la respuesta.	La respuesta se encuentra correcta y completa en su totalidad.
Pregunta 1	--	[0-4]%	[5-12]%	15%
Pregunta 2	--	[0-8]%	[9-18]%	20%
Pregunta 3	--	[0-10]%	[11-28]%	30%
Pregunta 4	--	[0-10]%	[11-33]%	35%

obs. En caso que un ítem no se encuentre o se encuentre incorrecto en su totalidad, el % equivale a 0%