

### **Taller 3**

### **ICI612 Electivo Profesional**

### **Text Mining**

**Eliana Providel Godoy – eliana.providel@uv.cl**

*03 de octubre, 2025*

**Información general:**

- Evaluación para ser respondida en grupo de dos estudiantes y de forma presencial.
  - Debe utilizar el dataset de la tarea de Análisis de Sentimiento y la biblioteca NELA <https://pypi.org/project/nela-features/>, Revisar jupyter disponible en Aula Virtual con ejemplo de uso de NELA
  - Recuerde referenciar el material externo que utiliza
  - Hora y fecha de Entrega: 03 de octubre, 2025 hasta las 13:30, por Aula Virtual.
  - Durante la evaluación en clases, habrá un porcentaje de la nota que será por preguntas que la profesora realizará a cada grupo de forma presencial (debe estar cada integrante del grupo).
1. (30%) Utilizando, al menos, entre 1000 y 1500 mensajes/reviews/post/... del dataset de la tarea de análisis de sentimiento y la biblioteca NELA, extraiga las siguientes características:
    - a. Complexity: 'avg\_wordlen', 'word\_count'
    - b. Bias: 'assertatives', 'factives', 'hedges', 'implicatives', 'positive\_opinion\_words', 'negative\_opinion\_words'
    - c. Extract\_event: 'num\_locations', 'num\_dates'
    - d. Extraiga dos características extras que usted considere relevante para su dataset, justifique su respuesta.
  2. (40%) Utilizando los resultados obtenidos en el punto 1, caracterice cada etiqueta de su dataset. La respuesta debe incluir dos gráficos.
  3. (30%) A cada grupo se le realizará una pregunta el día del control.

**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.
1 y 3	–	[1-9%]	[10-25%]	[30%]
2	--	[1-13%]	[14-35%]	[40%]

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