

Tarea 1: Algunos enlaces importantes para la tarea.

Ejemplo de limpieza de datos (recuerden que no es parte de la tarea, solo es por si les llega a servir):

https://github.com/mayraberrones94/FCFM/blob/master/P1_Limpieza_datos.ipynb

Otros ejemplos que pueden servirles:

<https://www.kaggle.com/code/ahmedterry/disneyland-reviews-nlp-sentiment-analysis/notebook>

<https://www.kaggle.com/code/ngyptr/python-nltk-sentiment-analysis>

<https://www.kaggle.com/code/tanulsingh077/twitter-sentiment-extraction-analysis-eda-and-model#About-this-Notebook>

Base de datos de Amazon food review (la que se uso en el ejemplo):

<https://www.kaggle.com/datasets/snap/amazon-fine-food-reviews?select=Reviews.csv>

Siguiendo el preprocesamiento de sus datos, realicen un análisis de sentimiento, haciendo comparación entre las diferentes librerías que se mencionan en clase. Se tomará en cuenta la discusión de resultados y su correcto análisis. Agregar en las conclusiones cual es la mejor librería que usaron y por qué.

Ejemplo en mi GitHub:

https://github.com/mayraberrones94/FCFM/blob/master/Semana_2_Analisis_de_sentimiento.ipynb

El trabajo será entregado de la siguiente manera:

- Archivo en formato PDF en la plataforma de Teams.
- Archivo PDF, documento de código e imágenes adjuntas, en su repositorio personal de GitHub.

Recuerden que en el archivo PDF deben de agregar en las referencias el URL de su GitHub personal y el URL de la página de donde sacaron su base de datos.

Se evaluará:

- En los resultados, que hayan utilizado al menos las 3 librerías que vimos en clase.
- En caso de que tengan que modificar los parámetros que se utilizaron en el ejemplo para la librería, explicar porque se hicieron y si fueron exitosos o no.
- En caso de utilizar el ultimo modelo (el de random forest) explicar sus resultados y sacar conclusiones sobre el uso de ese modelo.
- En caso de utilizar la misma base de datos que la clase anterior, atender a cualquier corrección que se haya echo en la revisión para evitar repetir errores.

Puntuación:

Las tareas individuales pueden llegar a tener hasta 6 puntos.

6 puntos - Excelente: Sobrepasa la expectativa del trabajo. Se agregaron elementos de interés a la tarea, como orientar el trabajo con elementos que sirvan para su proyecto de tesina.

5 – 4 puntos – Bueno: Cumple con lo que se espera de la tarea. Baja un poco si se tiene faltas de ortografía o inconsistencias en la discusión de resultados.

3 – 2 puntos – Regular: Faltan puntos que se comentaron para la tarea. Tiene algo que puede ser rescatable, pero falta mucho contenido.

1 punto – Malo: Se intento hacer algo, pero al final no tiene mucho sentido.

0 puntos: No entrego la tarea.

NOTA: Recuerden que la puntuación se evalúa a la mitad de los puntos que pudieron haber obtenido si lo entregan a destiempo por una semana. Después de la primera semana todos los trabajos se evalúan a 0.