

Tarea 1 - INFO279 - Tratamiento Automático del Lenguaje - 2024

Desafío 1: Construir un modelo de inteligencia artificial capaz de clasificar noticias de prensa según su categoría temática principal.

category: sociedad, salud, política, medioambiente, internacional, entretenimiento, economía, deportes, cultura, cienciatecnologia

Desafío 2: Geolocalizar eventos en noticias

Utilizando técnicas de tratamiento automático del lenguaje, el objetivo del desafío es estructurar la información de una noticia de la manera siguiente:

- **event**: principal evento descrito en la noticia
- **category**: categoría temática del evento
- **address**: dirección dónde ocurrió el evento en formato: calle número, comuna, país
- **latitud**: latitud del evento
- **longitud**: longitud del evento

Resultado: El formato del archivo CSV output que entregar contiene las columnas siguientes: id_news, event, category, address, latitud, longitud

Evaluación del desafío: Se evaluará su método en base a 100 noticias etiquetadas con las métricas siguientes:

- Category: Precision y Recall
- Longitud y Latitud: Exactitud (con una margen de error de 100 metros)

Pauta de evaluación

Nota	Descripción
7	El método implementado identifica el evento principal de la noticia con exactitud y logra geolocalizarlo con precision. Además categoriza la temática del evento.
6	El método implementado identifica el evento principal de la noticia y logra geolocalizarlo en un gran número de casos
5	El método implementado identifica el evento principal de la noticia, devuelve información de geolocalización aunque no es muy precisa.
4	El método implementado identifica un evento de la noticia aunque no es lo más preciso y devuelve información de geolocalización (no

	necesariamente de manera precisa)
--	-----------------------------------

- Notebook: se entregan por github
- archivos CSV: por discord