

Apellidos :

Nombre :

Sólo se corregirán los ejercicios escritos a **bolígrafo azul** o **bolígrafo negro**

Ejercicio 1. [1.5 puntos]

Una agencia de noticias desea optimizar el sistema de recepción de nueva información, sobre todo teniendo en cuenta la ingente cantidad de información que llega a través de determinadas redes, en particular Twitter.

Para ello nos pide diseñar un modelo que permita recibir un mensaje y dirigirlo directamente a una de las 4 secciones básicas en las que va a clasificar los mensajes. En concreto, se han elegido las secciones de Economía, Política, Deporte y Ocio para hacer un primer prototipo del sistema.

Como punto de partida, la agencia nos indica que dispone de los titulares de dichas secciones en los últimos 5 años (lo que supone unos 1000 titulares por cada una de las secciones). Así mismo, el trabajo realizado durante los últimos meses en la agencia de noticias, ha permitido clasificar manualmente unos 3000 tweets para cada una de las secciones.

Teniendo en cuenta este escenario, en este ejercicio nos piden:

- Explicar en qué consiste la clasificación automática de documentos y cómo la podemos aplicar en este contexto.
- Definir un algoritmo que nos permita llevar a cabo el entrenamiento a partir de los datos disponibles, y así mismo, aplicarlo para la clasificación de una nueva noticia recibida.

En concreto, nos piden que apliquemos el algoritmo Naive Bayes para la clasificación de documentos. Describe tanto el algoritmo de entrenamiento como el de clasificación para aplicarlo a este problema.

Ejercicio 2. [1.5 puntos]

Explica el algoritmo de retropropagación en redes neuronales.

Entre otras cuestiones, es importante cubrir los siguientes puntos:

- Para qué tipo de redes (con qué configuraciones) es adecuado utilizar este algoritmo.
- Qué fases principales incluye el algoritmo. Explica claramente los objetivos que persigue cada fase y ofrece una versión clara (a nivel de pseudocódigo) de cada fase.