

Normas de valoración del examen:

- La nota del examen representa el 60% de la valoración final de la asignatura (el 40% restante corresponde a las prácticas).
- Cada cuestión contestada correctamente vale 1 punto.
- SE PERMITE CALCULADORA

### Cuestiones:

1. Calcule la distancia de edición (usando coste 1 para todas las operaciones) de "terma" y "tierra". Indique las operaciones realizadas.
2. ¿En qué consiste el proceso de normalización temporal de las expresiones temporales?
3. Considere el siguiente etiquetado gramatical (POS tags) de una oración, donde la segunda columna corresponde al gold standard y la tercera a las etiquetas asignadas por un etiquetador que se está evaluando. Calcule la tasa de aciertos (exactitud del etiquetado) y las medidas de precisión, cobertura y medida-F correspondientes a las etiquetas NN y VBD.

Japan NNP NNP  
began VBD VBD  
the DT DT  
defence NN NN  
of IN IN  
their PRP\$ PRP\$  
Asian JJ JJ  
Cup NNP NN  
title NN NN  
with IN IN  
a DT DT  
lucky JJ JJ  
2-1 JJ JJ  
win VBP NN  
against IN IN  
Syria NNP NNP  
in IN IN  
a DT DT  
Group NNP NNP  
C NNP NNP  
championship NN NN

match NN NN  
on IN IN  
Friday NNP NNP  
...

4. ¿En qué consiste una matriz término-contexto? ¿En qué modelo de representación se utiliza?
5. Escribe el esquema general (descripción de los pasos principales) de los algoritmos de clustering iterativos de reubicación. Indica un algoritmo que pertenezca a esta familia.
6. Indica qué medidas de evaluación se utilizan habitualmente para comparar la salida de un sistema de clasificación automática con un conjunto de referencia o gold standard.