



**UNIVERSIDAD DE COSTA RICA
SISTEMAS DE ESTUDIO DE
POSTGRADO EN ESTADÍSTICA**



Examen parcial n° 2

Nombre: _____ Carné: _____

I. RESPUESTAS BREVES. (25 PUNTOS / 5 PTS CU).

1. Explique la diferencia entre un análisis por agrupamiento y un análisis por clasificación, brinde un ejemplo para cada caso.
2. Explique qué es el dendograma, cómo es que se juntan los casos, como se ve esto visualmente, y dónde es que se debería cortar para saber la cantidad de grupos.
3. Enuncie cada una de las partes de un árbol de decisión, e indique qué es cada una de los componentes anteriores.
4. ¿Cuál es la diferencia entre un árbol de decisión, un árbol de regresión y un árbol de clasificación? ¿Cuáles son los tipos de resultados esperados de cada uno?
5. ¿Cuál son los tipos de métodos de clasificación que existen? Explique en qué consiste cada uno de estos.

II. PREGUNTAS PRÁCTICAS. (75 pts / 25 pts CU). Los datos se obtienen de las librerías cluster.datasets y librería o el paquete dataset.

1. Utilizando el archivo de datos us.civil.war.battles, realice un análisis de clusters, bajo los dos enfoques que vimos. Interprete los resultados. ¿Qué puede concluir del presente análisis?

2. Realice el siguiente árbol de clasificación para los datos de Iris.
- Realice un árbol de clasificación.
 - Compare los resultados de la clasificación realizada con el método del análisis discriminante.

La referencia del archivo de datos se encuentra en el siguiente enlace:
<https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/iris.html>

3. Realice un análisis de clúster o clasificación para los datos de Iris.
- Determine para cada uno de los métodos la mejor clasificación
 - Compare los resultados de la clasificación realizada con el método de árboles y el discriminante.

La referencia del archivo de datos se encuentra en el siguiente enlace:
<https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/iris.html>

PUNTOS EXTRAS

4. Para el archivo de datos de Iris, comprueba que el método de ACP, discriminante, clasificación y árboles de decisión, permiten llegar a resultados equivalentes. (10 pts EXTRAS).

ESTA PREGUNTA ES LARGA Y LABORIOSA, SINO TIENEN CERTEZA, MEJOR NO LA HAGAN.

¡BUENA SUERTE!