

Primer examen parcial

Sistemas Inteligentes II

Jorge Alberto Jaramillo Garzón
Universidad de Caldas

- 1) (1.5) Implemente el procedimiento de **validación cruzada** de 10 particiones y úselo para determinar el desempeño de un clasificador bayesiano con distribución gaussiana sobre la base de datos *Iris* (considere solamente las **clases setosa** y **versicolor** pero use todas las características). Repita el procedimiento para un **clasificador geométrico** por mínimos cuadrados. ¿Qué diferencias encuentra? Documente los pasos, pruebas, errores y ajustes que realizó hasta obtener los resultados finales
- 2) (1.0) Implemente el clasificador K-NN sobre la base de datos *Iris* (usando las cuatro características y las tres clases) e investigue su desempeño mediante el método de *bootstrapping* a medida que se aumenta el número de vecinos empleado. Proponga una estrategia práctica de selección de k si este algoritmo se usara en un caso real con datos de otra naturaleza.
- 3) (1.0) Compare y contraste los tres clasificadores vistos en clase (geométrico, bayesiano y K-NN) según los siguientes criterios: ¿Qué suposiciones hace cada uno de los modelos? ¿Cuáles son sus requerimientos de entrenamiento en términos de tiempo de cómputo y memoria? ¿Cómo cambiará el desempeño de cada uno cuando se aumenta el número de dimensiones del espacio de características? Relacione cada clasificador con un ejemplo realista de aplicación (puede ser del contexto colombiano o de su experiencia personal).
- 4) (1.5) Sustentación oral en horario de clase.

Entregable: Un notebook de Jupyter con código, salidas y comentarios personales. Puede emplear las librerías que considere pertinentes para la implementación de los modelos de aprendizaje de máquina, *excepto para las estrategias de validación*.