Archivos adjuntos:

Archivo retail.csv

Archivo ionosphere.csv

Archivo Basketball.csv

Archivo apriori.py

**Ejercicio 1:** En el archivo retail.csv hay 18.000 tickets de compra anonimizados.

Cuestiones:

1. Obtener los conjuntos de dos, tres y cuatro elementos con mayor soporte.
2. Utilizando como soporte mínimo el obtenido en paso anterior para el conjunto de tres elementos, obtener todas las reglas de asociación con confianza de 0,8
3. Utilizando como soporte mínimo el obtenido en el primer paso para el conjunto de cuatro elementos, obtener todas las reglas de asociación con confianza de 0,5

**Ejercicio 2:** En el conjunto de datos Baskeball.csv se encuentran las siguientes columnas para diferentes jugadores:

* X1: altura en pies
* X2: peso en libras
* X3: porcentaje de tiros de campo exitosos
* X4: porcentaje de tiros libres exitosos
* X5: puntos promedio anotados por juego

Intentar predecir el porcentaje de tiros de campo exitosos para un jugador en base al resto de variables utilizando la regresión lineal y regresión LASSO con parámetros alpha de 0.01, 0.1 y 0.5.

Cuestiones:

1. ¿Cuál es el R2 en cada caso?
2. ¿Cuántas variables quedan disponibles al finalizar la regresión LASSO con cada valor de alpha?

Nota: En este caso no se ha de separar los datos en conjunto de entrenamiento y test. Se ha de trabajar siempre con el conjunto de datos completo.

**Ejercicio 3:** En el archivo ionosphere.csv hay datos de la ionosfera. Las 34 primeras columnas son valores numéricos y la ultima es una etiqueta de texto con ‘g’ cuando el resultado es bueno y ‘b’ cuando es malo. Importar este conjunto de datos y crear un modelo de clasificación utilizando clasificadores regresión logística, máquinas de vector soporte y k-nn con 3 y 5 vecinos.

Los datos se han de dividir en un conjunto de entrenamiento que contenga el 80% de los datos y otro test con el resto.

Cuestiones:

1. ¿Qué clasificador muestra una mayor precisión?
2. ¿Qué clasificador muestra un menor error de tipo I?
3. ¿Qué clasificador muestra un menor error de tipo II?