**Recuperatorio Segundo parcial AID Maestría en Data Mining agosto de 2017**

**Ejercicio 1**

En el archivo records masculinos.xls se encuentran los records masculinos de 55 países en diferentes actividades deportivas.

1. Realizar un análisis de cluster jerárquico explicando la clasificación obtenida y el método elegido para las actividades.
2. Idem para los países.
3. Explique la diferencia entre los métodos jerárquicos y no jerárquicos analizando en qué casos aplicaría cada uno de ellos.
4. Indique un modo de decidir el número de conglomerados para clusterización jerárquica y otro para clusterización no jerárquica.

**Ejercicio 2**

En la base **anticuerpos.xls** se registraron los datos de colesterol total, albumina, calcio y áurico de 188 pertenecientes a dos grupos definidos por la etiología de su patología.

Interesa saber si existen diferencias significativas entre los vectores medios de ambos grupos.

1. Establezca las hipótesis de interés. Realice el contraste correspondiente.
2. En caso de rechazar realice un análisis discriminante lineal o cuadrático según corresponda y explique cuál ha elegido y por qué.
3. Analice de dos maneras el poder discriminante de la regla.
4. A qué grupo asignaría un paciente con: Colesterol=240 ,Albumnia=39, Calcio=101 y aurico=49.

**Ejercicio 3**

Un ingeniero encargado de efectuar el control de calidad en una fábrica, desea estudiar la variabilidad existente, entre distintos lotes de productos manufacturados. La característica que debe investigar es el largo de cierto tipo de tornillos. Para ello cuenta con 50 cajones donde se encuentra la producción de un día. El ingeniero decide tomar aleatoriamente 5 cajones (lotes) y medir la longitud de 10 tornillos (en cm.) en cada uno de ellos. En la tabla siguiente se encuentran los datos tomados por el ingeniero:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lote 1 | Lote 2 | Lote 3 | Lote 4 | Lote 5 |
| 2.2 | 2.6 | 1.9 | 3.5 | 2.7 |
| 2.1 | 2.2 | 1.5 | 2.3 | 2.5 |
| 2.3 | 2.2 | 2.2 | 3.0 | 2.2 |
| 2.5 | 2.5 | 1.6 | 2.1 | 2.3 |
| 2.9 | 2.2 | 2.3 | 3.3 | 2.3 |
| 2.5 | 2.8 | 1.8 | 2.3 | 1.9 |
| 2.2 | 1.6 | 2.3 | 1.8 | 1.8 |
| 1.7 | 2.3 | 1.5 | 3.0 | 2.3 |
| 2.3 | 2.5 | 2.5 | 2.7 | 2.2 |
| 2.5 | 2.2 | 1.6 | 3.2 | 2.4 |

1. Especifique el modelo de análisis que va a utilizar y explique las razones de su elección.
2. Plantee la hipótesis de interés y proceda a probar la hipótesis nula establecida.
3. Verifique el cumplimiento de los supuestos del modelo. Es necesario realizar alguna transformación.
4. Estime la varianza del error experimental y concluya en el contexto del problema.