**Modulo 4. Parte 2. Clustering jerárquico.**

**Ejercicio** = Usar la matriz de distancias generada en el ejercicio de comparación de secuencias para crear una estructura de árbol (cluster jerárquico).

A screenshot of a computer

Description automatically generated

Como funciona el clustering jerárquico completo:

1. Encontrar la similitud mas alta para el primer cluster.
   1. Cluster 1 = CabbageMan y Corgi.
2. Actualizar la tabla remplazando a CabbageMan y Corgi por Cluster 1

|  |  |  |  |
| --- | --- | --- | --- |
|  | Berka | Yavru | Cluster 1 |
| Berka | 100 |  |  |
| Yavru | 9.114 | 100 |  |
| Cluster 1 | ? | ? | 100 |

* 1. Para seleccionar que valor usar para el nuevo cluster. Se escoje el valor mas bajo entre el nuevo cluster y cada genoma.

|  |  |  |
| --- | --- | --- |
|  | Berka | Yarvu |
| Corgi | **15.487** | 22.965 |
| Cabbageman | 15.677 | **21.185** |

* 1. Se escriben los valores en la tabla,

|  |  |  |  |
| --- | --- | --- | --- |
|  | Berka | Yavru | Cluster 1 |
| Berka | 100 |  |  |
| Yavru | 9.114 | 100 |  |
| Cluster 1 | 15.487 | 21.185 | 100 |

1. Se escoge la mayor similitud en la nueva tabla
   1. Cluster 2 = cluster 1 y Yavru
2. Se actualiza la tabla
   1. Escoger el valor menor entre {Berka y cluster 1 = 15.587} y {Berka y Yavru = 9.114}. La tabla actualizada es=

|  |  |  |
| --- | --- | --- |
|  | Berka | Cluster 2 |
| Berka | 100 |  |
| Cluster2 | 9.114 | 100 |

1. Ahora se arma el árbol:
   1. El primer cluster es Cabbageman y corgi

Cabbageman similitud = 93.2%

Corgi



* 1. El segundo cluster es Yavru y Cluster 1 (es Cabbageman y corgi)

Cabbageman



Corgi similitud = 15.5%



Yavru

* 1. El resultado final es Berka y Cluster 2 (Yavru y Cluster 1).

Cabbageman



Corgi



Yavru



Berka similitud = 9.1% \*Berka es la mas lejana

¿Esta estructura tiene sentido? Tenga en cuenta el resultado obtenido en VIRIDIC.