

Práctica 09

DOCENTE	CARRERA	CURSO
MSc. Vicente Machaca Arceda	Escuela Profesional de Ciencia de la Computación	Bioinformática

PRÁCTICA	TEMA	DURACIÓN
09	UPGMA	3 horas

1. Resultados del estudiante

- (a) Conocimientos en computación
- (b) Análisis de problemas.
- (c) Diseño y desarrollo de soluciones.
- (d) Trabajo individual y en equipo.
- (h) Uso de herramientas modernas.

2. Competencias de la práctica

- Implementa el algoritmo UPGMA para la construcción árboles filogenéticos.
- Utiliza herramientas como BioPython y ETE para la implementación y análisis de árboles filogenéticos.

3. Equipos y materiales

- Latex
- Python
- BioPython

4. Entregables

- Se debe elaborar un informe en Latex.
- El informe debe contener un enlace a github, pruebas y conclusiones.

5. Ejercicios

1. Instale las librerías BioPython, scikit-bio y ETE para el análisis de árboles filogenéticos. Además describa las librerías, en un párrafo de no mas de 5 líneas.
2. Evalúe el código a continuación y describa su funcionamiento.

```
1 from ete3 import PhyloTree, TreeStyle
2 from skbio import DistanceMatrix
3 from skbio.tree import nj
4
5 data = [[0, 8, 4, 6],
6         [8, 0, 8, 8],
7         [4, 8, 0, 6],
8         [6, 8, 6, 0]]
9 ids = list('abcd')
10
11 dm = DistanceMatrix(data, ids)
12 tree = nj(dm) # build a tree using neighbors joining algorithm
13 print(tree.ascii_art())
14
15 newick_str = nj(dm, result_constructor=str) # return newick format
16 print(newick_str)
17
18 t = PhyloTree(newick_str) # plot three using ETE
19 t.show()
```

3. Evalúe el código a continuación y describa su funcionamiento.

```
1 from ete3 import PhyloTree, TreeStyle
2 from skbio import DistanceMatrix
3 from skbio.tree import nj
4
5 fasta_txt = """
6 >seqA
7 MAEIPDETIQQFMALT---HNIQVQLSEFGLNEALNSYYASQTDDIKDRREEAH
8 >seqB
9 MAEIPDATIQQFMALTNVSHNIQVQY--EFGDLNEALNSYYAYQTDDQKDRREEAH
10 >seqC
11 MAEIPDATIQ---ALTNVSHNIQVQLSEFGLNEALNSYYASQTDDQKDRREEAH
12 >seqD
13 MAEAPDETIQQFMALTNVSHNIQVQLSEFGLNEAL-----REEAH
14 """
15
16 # Load a tree and link it to an alignment.
17 t = PhyloTree("((seqA,seqB),seqC),seqD);")
18 t.link_to_alignment(alignment=fasta_txt, alg_format="fasta")
19 t.show()
```

4. Implemente el algoritmo UPGMA. Este debe tomar como entrada una matriz de distancias y debe retornar el árbol filogenético en formato *newick*. Luego utilice la librería *ete3* para visualizar el árbol. Puede tomar como entrada la matriz de la pregunta 2.

6. Rúbricas

Rúbrica	Cumple	Cumple con obs.	No cumple
Implementación: Implementa de manera eficiente y correcta el algoritmo UPGMA [b, c].	10	5	0
Uso de herramientas: Aprende el correcto uso de las librerías scikit-bio y ETE durante el desarrollo de las preguntas 1, 2 y 3 [h].	5	2.5	0
Informe y presentación: Desarrolla un informe en Latex. Además, el alumno demuestra el trabajo en equipo (github) y dominio del tema durante la exposición [d].	5	2.5	0
Errores ortográficos: Por cada error ortográfico, se descontará 1 punto.	-	-	-