task_11.3

Machine Learning (WiSe 2025/2026)

Author: Suvansh Shukla
Matriculation No. 256245

Assignment 11 Task 3

Given:

- AACT, GAGG, GAGA, AACC
- Use Hamming distance
- generate HAC

Figuring out distance between the given sequences:

- AACT & GAGG = 3
- AACT & GAGA = 3
- AACT & AACC = 1

	AACT	GAGG	GAGA	AACC
AACT	0	3	3	1
GAGG	3	0	1	3
GAGA	3	1	0	3
AACC	1	3	3	0

Single Linkage

	AACT	GAGG	GAGA	AACC
AACT	0			
GAGG	3	0		
GAGA	3	1	0	
AACC	1	3	3	0

In single linkage we take the minimum value, which is 1 for (GAGA, GAGG) and (AACC, AACT)

We first merge GAGA and GAGG, resulting in this distance matrix:

	AACT	(GAGG, GAGA)	AACC
AACT	0		

	AACT	(GAGG, GAGA)	AACC
(GAGA, GAGG)	3	0	
AACC	1	3	0

Now we merge AACC and AACT, this leads to an updated distance matrix:

	AACT	(GAGG, GAGA)
(AACT, AACC)	0	
(GAGA, GAGG)	3	0

Now we just merge (AACT, AACC) & (GAGA, GAGG) for getting the final HAC.

Average Linkage

We start off with the same distance matrix:

	AACT	GAGG	GAGA	AACC
AACT	0			
GAGG	3	0		
GAGA	3	1	0	
AACC	1	3	3	0

Since we see the same sequence here again, i.e. 1 is the minimum distance for (GAGA, GAGG) and (AACC, AACT)

So we merge GAGG and GAGA first again giving us the same distance matrix as before

	AACT	(GAGG, GAGA)	AACC
AACT	0		
(GAGA, GAGG)	3	0	
AACC	1	3	0

Now we merge AACT & AACC, which gives us the same distance matrix as last time:

	AACT	(GAGG, GAGA)
(AACT, AACC)	0	
(GAGA, GAGG)	3	0

Now we just merge the two clusters to complete the HAC.

Complete linkage

We start off with the distance matrix:

	AACT	GAGG	GAGA	AACC
AACT	0			
GAGG	3	0		
GAGA	3	1	0	
AACC	1	3	3	0

Here too we see the same exact sequence merging GAGG & GAGA followed by AACC & AACT.

Ending by merging these two together.