Université Mohamed Khider-Biskra
Faculté des sciences exactes et des sciences de la nature et de la vie
Département d'informatique

M2 SIOD — Analyse des données séquentielles
26 Jan 2019 — 12:00-13:30

Examen de rattrapage

Barème: 1: 1.5 pts / **2:** 2 pts / **3:** 2pts / **4:** 2.5 pt / **5:** 4 pts / **6:** 6 pts / **7:** 2 pts

Soit les séquences d'ADN suivantes avec leur classes :

N°	Séquence	Classe
1	TTTTCAAGT	C1
2	ATTGGGGTA	C2
3	AAACGACCG	C2
4	GTTTTTGAC	C1
5	CGGCCATGA	C2
6	AATGGCAAA	C1

1. Représenter le d-plot de cet ensemble.
2. Calculer les durées moyennes passées dans chaque états pour les deux premières séquences.
3. Calculer l'entropie longitudinale des deux dernières séquences.
4. Sachant que le nombre de sous séquences de la troisième séquence est de 156. Calculer la turbulence de cette séquence.
5. En utilisant la méthode 3-KPPV avec la distance 2-spectrum, trouver la classe de la séquence $CGTGGTAGT$.
6. En utilisant l'algorithme AprioriAll et un support minimum de 50 %, trouver les motifs fréquents séquentiels.
7. En dédire les séquences maximales.

Bonne Chance

Dr A.Djeffal