DCD : Application de la découverte de motifs pour l'analyse de parties de Tennis de Table.

L'objectif de ce TP est d'appliquer des algorithmes de fouille de motifs afin d'analyser un partie de tennis de table. Il conviendra de s'interroger sur les types de motifs (langage) pouvant être extraits.

Modalité de rendu :

- Taille des groupes : 5 max
- A rendre via Tomuss avant le 05/12, 23:59 :
 - Un rapport sous la forme d'un notebook bien commenté où les choix devront être bien justifiés.
 - Le code associé

Les données :

Nous disposons d'un fichier décrivant une partie de Tennis de Table où les échanges ont été annotés.

https://perso.liris.cnrs.fr/marc.plantevit/ENS/DM/TT_game.csv

Objectif:

Nous souhaitons exploiter ces données afin de répondre à différentes questions permettant à un joueur de s'améliorer ou d'analyser le jeu de son adversaire :

- Quelles sont les combinaisons de jeu les plus efficaces (type de service, zones de jeu, etc.) ?
- Existe-t-il des comportements récurrents qui permettent à un joueur de gagner certains points ? Ou au contraire de les perdre.

Cette liste de questions n'est pas exhaustive, d'autres questions peuvent être imaginées. Vous devez à partir de ces données procéder à des extractions de motifs afin d'apporter des éléments de réponses à ces questions ou d'autres que vous aurez introduites. Il conviendra de se poser la question sur le type de motifs à extraire (plusieurs sont possibles). Des prétraitements seront nécessaires.

Liens (fortement conseillé de les parcourir avant de se lancer) :

Motifs séquentiels :

- https://github.com/Romathonat/MCTSExtent
- https://github.com/Romathonat/SegScout
- https://github.com/chuanconggao/PrefixSpan-py

 $Balance Span: \underline{https://github.com/guillaume-bosc/Balance Span}$

Subgroup discovery:

Pysubgroup: https://github.com/flemmerich/pysubgroup

Exceptional Model Mining:

https://github.com/MathynS/emm

Association rules:

- https://pypi.org/project/apyori/
- http://rasbt.github.io/mlxtend/