

Analyse de parcours de soin : Introduction, méthodologies, problèmes ouverts et apport des sciences des données

Vendredi 11 octobre 2024

De 9h30 à 17h00

CESP – Villejuif

Hôpital Paul Brousse Bat 15/16

16 avenue Paul Vaillant Couturier - 94807 Villejuif

10 :00 – 12 :00 :

- **Parcours de soins dans le SNDS : Application aux maladies infectieuses et difficultés rencontrées**

Watier Laurence, CESP

Après une brève présentation des données du Système National des Données de Santé, l'analyse de séquence utilisée pour étudier les parcours de soins sera exposée. Initialement utilisée en sciences sociales, l'analyse de séquences a été plus récemment appliquée à l'étude des parcours de soins. Cette méthode permet de décrire la dimension temporelle de la consommation de soins de santé et d'identifier différents modèles de soins et profils de patients. L'analyse des séquences décrit des données longitudinales par l'analyse d'ensembles de données séquentielles (ensembles de séquences). Chaque séquence est divisée en unités temporelles et, pour chaque unité, des états sont pris en compte. Après l'analyse de séquence, les séquences sont regroupées en différents groupes/clusters à l'aide de méthodes de classification. Enfin, deux applications seront présentées pointant les difficultés rencontrées, en particulier lorsque la taille de l'échantillon est importante.

- **Extraction de modèles formels capables de reproduire le caractère probabiliste des observations d'un système**

Paolo Ballarini, MICS, CentraleSupélec

Le "process mining" vise à dériver des modèles formels capables de reproduire le comportement d'un processus organisationnel donné en analysant les exécutions observées dans un journal d'événements (event log). Les éléments d'un event log sont des séquences finies (appelées traces ou mots) d'actions. De nombreux algorithmes de process mining ont été introduits qui permettent d'extraire un modèle (généralement sous forme de réseau de Petri) visant à reproduire, aussi précisément que possible, le langage d'un event log considéré. Cependant, étant donné que des exécutions identiques peuvent être observées plusieurs fois, les traces d'un journal

d'événements sont associées à une fréquence (d'occurrence) et, par conséquent, un journal d'événements produit intrinsèquement ce qu'on appelle un langage stochastique, c.a.d. un langage dont chaque mot (comportement) est associé à une probabilité de se produire dans le système observé. Du coup l'extension stochastique du process mining se pose comme objectif de dériver des modèles stochastiques (réseaux de Petri stochastiques) capables de reproduire la vraisemblance des exécutions observées. Dans cet exposé, je vais donner un aperçu du problème du process mining stochastique, en soulignant les principales difficultés qu'il comporte et en présentant certaines méthodes proposées dans la littérature jusqu'à présent.

13 :30 – 15 :30 :

- **Chaînage aux données du SNDS dans le cadre d'un essai clinique : Application à un parcours de soins en oncogériatrie**

Arnaud Bayle CESP Oncostat U1018

Arnaud Pages CESP Oncostat U1018

Après une présentation du protocole de l'essai clinique, les enjeux du chaînage aux données du Système National des Données de Santé (SNDS) seront présentés avec les difficultés rencontrées et les pistes de solution.

- **Conformance checking of (time aware) process models**

Thomas Chatain, LMF, ENS Paris-Saclay

Conformance checking techniques assess the suitability of a process model in representing an underlying process, observed through a collection of real executions. An important problem for conformance checking is to align a log trace with a model, that is, to find the minimal changes necessary to correct a new observation to conform to a process model. We will also introduce conformance checking for timed models, that is, process models that consider both the sequence of events in a process as well as the timestamps at which each event is recorded.

Inscription :

<https://forms.office.com/e/W8h2xbPHHS>

