



Sujet de thèse CIFRE

Sujet : Modélisation et analyse de parcours de soins à l'aide d'outils d'intelligence artificielle

Encadrement académique : Mines Saint-Etienne (EMSE)

Localisation principale : Sud de la France

Laboratoire : UMR CNRS 6158 LIMOS

Encadrants et directeurs de thèse : EMSE : Vincent Augusto (directeur), Benjamin Dalmas (co-encadrant)

Date de démarrage : 1^{er} octobre 2019

Salaire brut annuel approximatif : 23 484 euros (1 957 euros/mois)

1 Contexte, objectifs et méthodologie de recherche

1.1 Contexte

La coordination des parcours de soins est un enjeu majeur pour optimiser les dépenses de santé et améliorer la qualité des soins. Dans cette optique, la société partenaire développe une expertise autour des activités suivantes :

- Accompagner et assister les équipes médicales dans la mise en place de leurs dispositifs de coordination.
- Renforcer la collaboration ville-hôpital en organisant une prise en charge transversale.
- Former les équipes de soins aux nouveaux métiers de la coordination des parcours de soins.
- Concevoir, développer et maintenir des solutions e-santé dédiées à la coordination.

Après plusieurs années d'activité, la société possède une base de données conséquente et souhaiterait développer de nouveaux modules intelligents autour de sa solution logicielle. En particulier, il serait intéressant de mieux prévoir les besoins de nouveaux patients s'il était possible de prédire la complexité de sa coordination. Par ailleurs, les données disponibles permettent de réaliser de nouvelles analyses descriptives telle la découverte de parcours caractéristiques, ou le benchmark d'instituts partenaires de la société.

1.2 Objectif

L'objectif principal de cette thèse vise à proposer plusieurs analyses des données de parcours de soins à partir des données dont dispose la société. Plus précisément, ce projet de thèse vise à :

1. Evaluer les performances d'un réseau coordonné par la société vs un réseau non coordonné à partir des données du Système National des Données de Santé.
2. Découvrir des parcours de soins caractéristiques d'une pathologie et les simuler.
3. Classer les instituts selon un ou plusieurs critères de performance et proposer des approches d'amélioration.
4. Modéliser les interactions entre les acteurs des parcours de soins.

Plusieurs verrous scientifiques devront être levés afin d'atteindre ces objectifs :

- Développer une approche adaptée de fouille de processus (process mining) sur les données disponibles. Plusieurs extensions vers la simulation et la comparaison de graphes devront être développées spécifiquement pour notre problématique.
- Proposer un modèle adapté aux communications et interactions dans un réseau de prise en charge.

2 Méthodologie de recherche

2.1 Caractérisation de la complexité de coordination d'un parcours de soins

Afin d'atteindre les objectifs de l'étude et de couvrir l'ensemble des points d'attention identifiés dans la section précédente, l'équipe Ingénierie des Systèmes de Soins et des Services de Santé (I4S) du Centre Ingénierie et Santé (CIS) de l'école des Mines de Saint-Etienne (EMSE) propose une approche de modélisation mathématique adaptée au problème. Cette approche vise à déterminer une ou plusieurs fonctions quantitatives adaptées permettant de quantifier la complexité de la coordination d'un parcours de prise en charge. Cette caractérisation permettra de définir plusieurs variables additionnelles, construites à partir des données existantes, caractérisant la complexité d'un patient à différents instants de sa prise en charge.

Par exemple, la complexité de coordination pourra être définie en fonction du nombre d'événements impliquant la coordinatrice d'un réseau, ou du nombre d'appels échangés entre la coordinatrice et le patient.

2.3 Découverte, simulation et comparaison de parcours de soins

Cette analyse permet d'appliquer une approche de type découverte de processus (process mining) sur un ensemble de données. Cette approche a été largement utilisée dans le domaine de la santé, notamment sur la base du PMSI (Programme de Médicalisation des Systèmes d'Information (Prodel et al. 2018) et offre la possibilité de découvrir des caractéristiques « cachées » dans des masses de données. Le résultat d'une telle approche est un graphe permettant de visualiser les parcours les plus caractéristique d'une population à partir d'un journal d'événements.

Les algorithmes classiques de process mining seront appliqués sur les données disponibles. Plusieurs modèles seront proposés :

- Par zone géographique (représentation de parcours suivis sur plusieurs instituts/zones géographiques).
- Par type d'activité (activités de soins seulement, activités de coordination seulement, toutes activités)

Selon les résultats, une phase de préparation des données pourra être appliquée par profilage des patients (Song et al. 2008) : il s'agit d'appliquer une méthode de classification non supervisée afin de déterminer des groupes de patients homogènes, et qui ont donc de plus grandes chances de donner lieu à des modèles de parcours intelligibles.

Après la découverte des graphes, deux approches seront explorées :

1. La simulation de flux, permettant de tester des scénarios de prise en charge ou d'innovation.
2. La comparaison quantitative de graphes de parcours. Pour cette étape, nous pourrions procéder à un appariement des données de la société avec le SNDS afin de comparer des cohortes prises en charge vs des cohortes sans prise en charge.

2.4 Benchmark des instituts de coordination

A partir des modèles déterminés par zone géographique, nous pouvons extraire plusieurs indicateurs de performance (à préciser). Un classement des instituts pourra être proposé, permettant d'établir un premier état des lieux sur la performance globale des réseaux suivis sur l'ensemble du territoire. Cette étude préliminaire pourra donner lieu en perspective au développement d'approches DEA (Data Envelopment Analysis) (Ozcan 2014) visant à proposer des pistes d'amélioration aux instituts les moins bien classés.

3 Profil recherché

Le candidat devra être titulaire d'un Master 2 Recherche ou équivalent et posséder des compétences en informatique, sciences des données, génie industriel et recherche opérationnelle. Une connaissance du domaine de la santé sera un plus.

Envoyer CV, lettre de motivation, lettres de recommandation et relevés de notes des 3 dernières années à Vincent AUGUSTO – augusto@emse.fr

4 Références

(Augusto and Xie 2014) V. Augusto and X. Xie, A modelling and simulation framework for health care systems, IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, Institute of Electrical and Electronics Engineers (IEEE), 2014, 44 (1):30-46, 2014.

(Augusto et al. 2015) Augusto V., O. Rejeb, X. Xie, S. Aloui, L. Perrier, P. Biron and T. Durand, 2015 : "Performance Evaluation of Health Information Systems using ARIS modeling and discrete-event simulation", Proceedings of the 2015 Winter Simulation Conference, Huntington Beach, CA, USA.

(Hooijenga et al. 2018) D. Hooijenga, R. Phan, V. Augusto, X. Xie and A. Redjaline, "Discriminant analysis and feature selection for emergency department readmission prediction," *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, Bangalore, India, 2018, pp. 836-842. doi: 10.1109/SSCI.2018.8628938

(De Oliveira et al. 2018) H. De Oliveira, M. Prodel and V. Augusto, "Binary Classification on French Hospital Data: Benchmark of 7 Machine Learning Algorithms," 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 2018, pp. 1743-1748. doi: 10.1109/SMC.2018.00301

(Feillet et al. 2014) Dominique Feillet, Thierry Garaix, Fabien Lehuédé, Olivier Péton, Dominique Quadri. "A new consistent vehicle routing problem for the transportation of people with disabilities", Networks, Volume 63 3:211-224, 2014.

(Liu et al. 2014) Ran Liu, Xiaolan Xie, Thierry Garaix. "Hybridization of Tabu Search with Feasible and Infeasible Local Searches for Periodic Home Health Care Logistics", Omega, 47:17-32, 2014.

(Ozcan 2014) Ozcan, Yasar. (2014). Performance Measurement Using Data Envelopment Analysis (DEA). 10.1007/978-1-4899-7472-3_2.

(Prodel et al. 2018) M. Prodel, V. Augusto, B. Jouaneton, L. Lamarsalle and X. Xie, "Optimal Process Mining for Large and Complex Event Logs," in IEEE Transactions on Automation Science and Engineering, vol. 15, no. 3, pp. 1309-1325, July 2018. doi: 10.1109/TASE.2017.2784436

(Siqiao et al. 2015) Siqiao Li, Na Geng, Xiaolan Xie. "Radiation Queue: Meeting Patient Waiting Time Targets", IEEE Robotics and Automation Magazine, 22/2, 51-63, 2015.

(Song et al. 2008) Song, Minseok & W. Günther, Christian & Aalst, Wil M. P.. (2008). Trace Clustering in Process Mining. Lecture Notes in Business Information Processing. 17. 109-120. 10.1007/978-3-642-00328-8_11.