

Méthodes d'Appariement Optimal pour l'analyse des parcours de soins

Romain Demeulemeester^{a,b,c}, Fanny Crozes^{b,c}, Nadège Costa^{b,c}, Philippe Saint-Pierre^{d,e}, Nicolas Savy^{d,e}

^a Université Toulouse III Paul Sabatier; Toulouse, France.

^b Unité d'Evaluation Médico-Economique, Département d'Information Médicale, CHU de Toulouse ; Toulouse, France.

^c INSERM (UMR 1027), Faculté de médecine de Toulouse ; Toulouse, France.

^d Institut de Mathématiques de Toulouse ; UMR5219; Toulouse, France.

^e Université de Toulouse; CNRS; Toulouse, France.

Dans le cadre de l'analyse des parcours de soins, la capacité à construire des clusters de trajectoires permettant l'identification de typologies particulières, sans connaissances préalables, constitue actuellement un enjeu important dans le domaine de la classification non supervisée.

Les méthodes de classifications non supervisées habituelles reposent cependant sur la capacité d'exprimer une mesure de dissimilarité entre les objets à l'étude, ce qui, dans le contexte de l'analyse de trajectoires de soins, représente un exercice non trivial.

En 1986, Andrew Abbott introduit pour la première fois la méthode d'Appariement Optimal en sciences sociales via l'étude de processus historiques [1]. Cette méthode adapte la distance de Levenshtein, permettant d'établir une mesure de la différence entre deux chaînes de caractères, au cadre de séquences d'états successifs. Elle repose sur une minimisation du coût nécessaire à la transformation d'une séquence en une autre via trois opérations élémentaires qui sont l'insertion, la suppression et la substitution d'un élément de la séquence, chacune étant affectée d'un coût élémentaire.

Les dissimilarités ainsi obtenues sont ensuite utilisées pour mettre en place des méthodes de Classification Ascendante Hiérarchique visant à constituer des clusters de trajectoires de soins. S'en suit alors une recherche de typologies particulières au sein des groupes construits.

Les méthodes explicitées seront illustrées par les travaux effectués par une étudiante Master 2 Recherche en Epidémiologie clinique (Fanny Crozes) au cours de son stage intitulé « Analyse socio-économique et économique du parcours de soins hospitalier des patients atteints d'AVC ischémique ».

[1] Andrew Abbott and John Forrest. Optimal Matching Methods for Historical Sequences. *Journal of Interdisciplinary History*, Vol. 16, No. 3 (Winter, 1986), pp. 471-494. The MIT Press.