

E07 : Apprentissage profond appliqué aux parcours de soin de patientes atteintes de cancer du sein : ouvrir la boite noire

Titre

Français : Apprentissage profond appliqué aux parcours de soin de patientes atteintes de cancer du sein : ouvrir la boite noire
Anglais : Deep learning applied to care pathways: open the black box

Auteurs

D Kannoun (1), C Bic (2), P Rinder (2), P Hornus (2), T Janssoone (2)
(1) Gynecologie, Clinique Pasteur, 45, avenue du Lombez , 31076, Toulouse, France
(2) Data Science, Sêmeia, 9 cour des petites écuries, 75010, Paris, France

Responsable de la présentation

Nom : Janssoone
Prénom : Thomas
Adresse professionnelle : 9 cour des petites écuries
Code postal : 75010
Ville : Paris
Pays : France
Newsletter :

Mots clés

Français : Apprentissage profondSNIIRAMexplications de décision
Anglais : Deep learningSNIIRAMexplanations of decision

Spécialité

Principale : Gynécologie
Secondaire : Oncologie - Fertilité

Texte

La persistance représente le respect de la durée d'un traitement médicamenteux jusqu'à son terme et ce, sans interruption de celui-ci. Le fait de ne pas respecter le plan de traitement est en réalité un problème majeur, car cela compromet gravement l'efficacité de thérapie à long terme et augmente le coût des services de santé[1]. Nous présentons ici nos travaux sur la modélisation de la consommation de médicaments par les patientes dans les traitements du cancer du sein afin de pouvoir prédire les moments à risque de non-persistance. À partir des données de remboursement du système de santé français (SNIIRAM), nous reconstruisons les parcours de soins des patients. Le SNIIRAM est l'une des plus grandes bases structurées de données de la santé au monde. L'utilisation de ces données massives permet l'application de modèles complexes et la détection de signaux faibles. Les données utiles sont, par exemple, les hospitalisations, les achats de médicaments ou les informations contextuelles sur le patient (âge, services gouvernementaux, informations géographiques,...) [2].

Ces dernières années, des algorithmes d'apprentissage automatique ont été développés afin d'analyser ces grandes bases de données. Nous comparons ici différentes approches de modèles computationnels pour l'analyse des facteurs de non-persistance. La cohorte de l'étude comprend les femmes qui, entre 2013 et 2015, répondent aux critères suivants : diagnostiquée d'un cancer du sein, ayant acheté une des molécules étudiées (Anastrozole, Exemestane, Letrozole, et Tamoxifene). Nous avons retravaillé les données brutes (transactions pharmaceutiques et hospitalisations) pour montrer les différentes phases du traitement. Une phase est une période d'absorption continue d'une molécule ou d'hospitalisations pour chimiothérapie ou radiothérapie, reflétant ainsi le parcours du patient. Une phase du traitement est considérée comme censurée par l'un des arrêts légitimes (décès, changement de traitement, problème cardiaque grave ou début de soins palliatifs) si cet événement survient moins de deux mois après la date de la dernière dose théorique (stock acheté et hospitalisations).

Nous avons comparé différents formalisme et algorithmes pour prédire un arrêt illégitime. Une analyse statistique basée sur des modèles de survie montre l'évolution du risque dans le temps. Un modèle de Cox permet également d'évaluer l'influence de chaque variables sur ces arrêts.
Ensuite, différentes méthodes d'apprentissage automatique (régression logistique, arbre de décision, gradient boosting, MLP) pour la prédiction d'un arrêt illégitime après une durée spécifique (3, 6 et 12 mois) à compter du début de la phase permettent de prédirent avec une efficacité (AUC) de 0.7 et de repérer les patients à risques avec 2 fois plus de précision.
Enfin, les dernières méthodes issues de l'apprentissage profond (LSTM, GRU) vont prédire le risque à chaque transaction, avec une AUC de 0.8 et

une précision 4 fois plus grande. Nous soulignons l'inconvénient de ces méthodes dans la difficulté à expliciter leur processus de décisions et soulignons nos premiers résultats sur leur interprétabilité. Nous montrons ainsi les intérêts et limites actuels de ces modèles.

Références :
[1] Krueger, Berger, et Felkey, 2005. Medication adherence and persistence: a comprehensive review. Advances in therapy , 22(4):313–356
[2] Tuppin, De Roquefeuil, Weill, et al. (2010) Revue d'epidemiologie et de sante publique, 58(4):286–290