

E03 : Test et validation du projet RUBY : Utilisation d'algorithme d'intelligence artificielle pour la structuration des données des comptes rendus médicaux de patientes atteintes d'un cancer du sein

Titre

- Français :** Test et validation du projet RUBY : Utilisation d'algorithme d'intelligence artificielle pour la structuration des données des comptes rendus médicaux de patientes atteintes d'un cancer du sein
- Anglais :** Test and validation of RUBY: use of artificial intelligence algorithm for the structuration of health record data for patient with breast cancer

Auteurs

R SCHIAPPA (1), G UZBELGER (2), B THAMPHYA (1), A FABRE (2), S TOLEDANO (2), J GAL (1), O SCHNEIDER (3), C GARCIA (2), S BENAVIDES (2), J GODFROY (3), J HAUDEBOURG (4), C BAILLEUX (5), JM FERRERO (5), O HUMBERT (6), E CHAMOREY (1), E BARRANGER (5)

(1) DEBDS, Centre Antoine Lacassagne, 33 avenue de valombrose, 06189, Nice, France

(2) Artificial Intelligence - Advanced Analytics Solutions, IBM, 17 Avenue de L'Europe, 92275, Bois-Colombes, France

(3) DSI, Centre Antoine Lacassagne, 33 avenue de valombrose, 06189, Nice, France

(4) ACP, Centre Antoine Lacassagne, 33 avenue de valombrose, 06189, Nice, France

(5) Oncologie-Sénologie, Centre Antoine Lacassagne, 33 avenue de valombrose, 06189, Nice, France

(6) Médecine Nuclaire, Centre Antoine Lacassagne, 33 avenue de valombrose, 06189, Nice, France

Responsable de la présentation

Nom : SCHIAPPA

Prénom : Renaud

Adresse professionnelle : 33 avenue de valombrose

Code postal : 06189

Ville : Nice

Pays : France

Téléphone : 0492031679

Adresse email : renaud.schiappa@nice.unicancer.fr

Newsletter :

Mots clés

Français : intelligence artificielle, machine Learning, deep Learning, cancer du sein

Anglais : artificial intelligence, machine Learning, deep Learning, breast cancer

Spécialité

Principale : Autres

Texte

Contexte

La structuration des données de santé dans les centres hospitaliers est une étape indispensable pour leur exploitation, cependant elle demeure très chronophage. Avec plus de 3 millions de comptes rendus (CR) et environ 200.000 nouveaux CR par an au Centre Antoine Lacassagne (CAL), la structuration manuelle rétrospective n'est pas une solution envisageable car elle mobiliserait beaucoup trop de ressources humaines. RUBY est un programme informatique développée au CAL en collaboration avec IBM qui permet la structuration automatique des données à l'aide d'algorithmes d'IA. L'objectif de ce travail est de tester rétrospectivement RUBY sur les patientes atteintes de cancer du sein au CAL.

Objectifs

Le programme RUBY visait à développer des algorithmes d'IA (réseaux de neurones convolutifs) capables de structurer automatiquement les CR des patientes atteintes d'un cancer du sein. La 1ère étape de ce projet était de développer des algorithmes d'IA sur une cohorte d'apprentissage puis de valider ces algorithmes sur une cohorte de test. La seconde étape, que nous présentons ici, est d'évaluer la précision et l'utilité de cet outil en comparant les résultats de la structuration des données par RUBY à une structuration manuelle.

Méthodes

RUBY a été appliqué aux 227 patientes de l'étude clinique promue au CAL : IODINE BREAST (IB). Les données d'IB ont été saisies par des ARC puis ont été monitorées afin de s'assurer de la qualité des données. La précision de RUBY a été évaluée en comparant les données saisies manuellement avec celles obtenues avec l'algorithme. RUBY a ensuite été appliqué sur la cohorte CHIRANA (extraite du PMSI) contenant 5702 patientes opérées d'un cancer du sein au CAL. L'objectif était de montrer la faisabilité et la rapidité de RUBY sur une base de données volumineuse.

Résultats

Au total, 19 variables ont été testées sur la base IB. L'algorithme RUBY atteint une précision >95% pour l'identification des comptes rendus de chirurgies, le type de chirurgie et le côté opéré ; une précision supérieure >90% pour les variables identification des compte rendus, SBR, HER2 amplification, expression œstrogène, date du compte rendu ; une précision >85% pour les variables côté de la tumeur, expression progestérone, curage axillaire, recherche de ganglion sentinelle, date de chirurgie. La précision est <85% pour les variables taille de la tumeur et type histologique. Quatre minutes ont été nécessaires pour la structuration des 19 variables des 227 patientes d'IB. Pour CHIRANA, 25 variables ont été structurées en 1h17 pour les 5702 patientes, à partir des comptes rendu de chirurgie et d'anatomopathologie.

Discussions

Les 1ères applications de RUBY sont très encourageantes. Une nouvelle phase d'apprentissage est nécessaire afin d'améliorer les performances notamment sur les tailles tumorales et les types histologiques. Les performances sont meilleures lorsque les comptes rendus testés sont contemporains de ceux de la période d'apprentissage de l'application. En effet, le vocabulaire et le format des comptes rendus ne sont pas stables dans le temps (nouveaux médecins, nouvelles assistantes médicales).

Conclusion

Cette collaboration avec IBM a permis de faire un grand pas en avant dans la structuration automatique des données du cancer du sein. RUBY permet un gain de temps considérable et sera un véritable succès lorsque toutes les précisions seront >95%. Des modèles pour d'autres pathologies sont en cours de développement au CAL.