

# **Prise en compte de données séquentielles hétérogènes dans l'apprentissage profond : application aux données de soins intensifs**

Mamadou Ben Hamidou Cissoko\*, Vincent Castelain\*\*, Nicolas Lachiche\*

\* ICube, Université de Strasbourg,

\*\* Hôpitaux Universitaires de Strasbourg

## **1 Problématique**

L'adoption généralisée des dossiers médicaux électroniques (DME) augmente l'intérêt pour les algorithmes d'apprentissage machine (ML) dans le domaine médical. Les données extraites du DME présentent souvent un grand nombre de valeurs manquantes pour les variables cliniques car les prélèvements de ces variables sont effectués à intervalles de temps irréguliers. Cela s'explique souvent par manque de collecte ou selon l'état sous-jacent du patient. Afin de résoudre ce problème, plusieurs stratégies ont été utilisées allant des méthodes d'imputation basées sur les noyaux, des procédures statistiques de base pour l'imputation, telles que le zéro, la moyenne à des stratégies permettant de directement modéliser les observations contenant des valeurs manquantes. (Rajkomar et al., 2018; Song et al., 2018) ont proposé une méthode générale pour représenter ces séries temporelles d'événements multivariés irréguliers en séries temporelles régulières non biaisées en agrégeant les mesures en intervalles de temps discrets pour chaque variable médicale à chaque épisode de soins pour créer des séries temporelles multivariées avec un intervalle de temps régulier.

## **2 Approche proposée**

Nous proposons un réseau neuronal multimodal basée sur LSTM capable de gérer la nature des irrégularités présentes dans les données DME, tant au niveau temporel qu'au niveau des caractéristiques mesurées. Il prends en entrée les données des patients et les indicateurs binaires de l'absence de chaque variable à chaque pas de temps (les données sont modélisées par heure). Nous montrons l'efficacité de l'approche proposée sur deux tâches médicales de référence (mortalité et durée de séjour) utilisant des données cliniques réelles dérivées des DME (MIMIC-III) (Johnson et al., 2016) pour des patients ayant séjourné dans les différentes unités des soins intensifs.