

# Fiche de poste pour un stage de Master 2 en Informatique

## Sujet du stage

Création de systèmes d'aide à l'annotation temps réel des dossiers patients hospitaliers à partir de données multimodales.

## Lieu de stage et contexte

Le stage se déroulera sur le campus universitaire de Luminy au Laboratoire d'Informatique et Systèmes (LIS) à Marseille.

Le stagiaire travaillera dans le cadre d'une collaboration interdisciplinaire autour des thématiques OneHealth rassemblant des membres de l'IRD ESPACE-DEV, et de la faculté des Sciences Médicales et Paramédicales de Marseille, et du LIS. Le stage sera financé par l'institut Laennec, un institut interdisciplinaire centré sur les applications d'IA à la santé.

## Contexte et problématique

En plus d'une meilleure utilisation des données existantes, l'intelligence artificielle promet une aide pour la génération de données de meilleure qualité tout en allégeant la charge de collecte de données pour les professionnels de santé. A ce titre, les données de diagnostics tirés de la Classification Internationale des Maladies (CIM), collectées à l'échelle mondiale à des fins de facturation et épidémiologiques, représentent, une source précieuse d'information structurée et un enjeu particulier. En effet, l'annotation par des labels CIM, qui doit être réalisée pour chaque séjour de soin hospitalier, est une tâche complexe, chronophage pour le personnel médical, et conduit à la production de données de faible qualité.

Le projet s'inscrit dans la continuité de travaux déjà effectués visant à dégager des principes généraux pour la construction et l'évaluation de systèmes d'aide au codage CIM.

## Sujet

Nos travaux précédents, utilisant une architecture de réseau de neurones séquentiel simple, ont montré que les prescriptions médicamenteuses représentaient une source d'information structurée de valeur et sous utilisée pour la création de système d'aide au codage CIM. Un premier volet du stage sera de développer des algorithmes d'apprentissage plus performants pour une meilleure exploitation des prescriptions médicamenteuses à l'aide de réseaux récurrents.

Comme mentionné précédemment, l'annotation CIM est une tâche complexe régie par de nombreuses règles, comme des exclusions, induisant des corrélations entre les occurrences de différents labels. Un second volet du stage consistera donc à développer des algorithmes exploitant ces corrélations et permettant la prise en compte des saisies utilisateur en temps réels à l'aide de modèles type autoencoders.

Le troisième volet du stage consistera à combiner ces deux approches, en y adjoignant les informations issues du texte libre et des hospitalisations passées pour la création d'un système multimodal temps réel.

## Objectifs

L'objectif de ce stage de Master 2 recherche sont :

- la recherche d'une architecture de réseaux de neurones efficace pour l'exploitation des prescriptions médicamenteuses pour la prédiction de codes CIM
- la fusion multimodale d'informations provenant des prescriptions médicamenteuses, des hospitalisations passées, et du texte libre pour la recommandation de codes CIM
- l'exploitation des corrélations inter-labels pour l'adaptation des recommandations en temps réel aux saisies utilisateur

## Profil du stagiaire, compétences requises et expérience :

Étudiant/e de Master en informatique ayant

- des compétences avancées en programmation (conception et implémentation) en Python (préférez), R ou Julia, et une culture des pratiques utiles à la reproductibilité (environnements virtuels, unit testing, intégration continue)
- des connaissances théoriques et pratiques sur les méthodes d'apprentissage par réseaux de neurones
- des connaissances générales dans les domaines de l'analyse des données et l'intelligence artificielle
- une appétence pour les problématiques liées à la santé et la qualité des données

## Modalités pratique du stage :

Le stage se déroulera au LIS site de Luminy, Marseille pour une durée de 5 mois.

Début du stage souhaité à partir de début février/mars jusqu'à fin juillet.

Temps de travail hebdomadaire : 35 heures.

Indemnités : Le stage est gratifié selon la législation en vigueur.

Le/la stagiaire sera en contact hebdomadaire avec ses encadrants.

## Candidature

Merci d'envoyer votre candidature (CV, deux derniers relevés de notes de Master 1) aux trois adresses suivantes : [quentin.marcou@univ-amu.fr](mailto:quentin.marcou@univ-amu.fr), [noel.novelli@lis-lab.fr](mailto:noel.novelli@lis-lab.fr) et [laure.berthel@ird.fr](mailto:laure.berthel@ird.fr)

Les candidatures sont gérées au fil de l'eau. Vous serez prévenu rapidement par mail de la recevabilité de votre candidature et si vous êtes invité à un premier entretien.