

STAGIAIRE EN ANALYSE DE DONNEES – H/F

Sujet

Aide à la décision pour la classification de séries temporelles de microbiologie.

Mission

Le sujet du stage porte sur la classification dite précoce de séries temporelles. Il existe en effet de nombreuses applications où des données sont collectées séquentiellement au cours du temps, et où l'on attend un résultat issu de leur traitement le plus tôt possible, quand bien même les données sont incomplètes. C'est le cas par exemple dans le milieu hospitalier, où établir un diagnostic médical est une tâche complexe qui se doit d'être réalisée le plus tôt possible pour une prise en charge optimale du patient. Pour ce faire, il faut être capable d'analyser les données collectées au cours du temps sur le patient. Il faut aussi effectuer le meilleur compromis entre retarder une prise de décision en faveur de données complémentaires ou privilégier la rapidité de la décision au détriment du doute.

Le stage portera sur ces deux problématiques : être capable de **classifier des séries temporelles** parfois incomplètes et **trouver l'instant optimal** (le plus tôt) à partir duquel la classification est fiable. C'est au sein de l'Unité Innovation de bioMérieux que vous contribuerez à la définition de méthodes et au développement d'algorithmes pour ces problèmes. Vous travaillerez sur des séries temporelles de microbiologie représentant des micro-organismes en phase de croissance et placés au contact de différents antibiotiques. La classification portera sur l'étude de la résistance des différents micro-organismes aux antibiotiques. Pour répondre aux problèmes posés **vous étudierez des algorithmes d'apprentissage par renforcement (Reinforcement Learning), champ de recherche actif dans le domaine de l'intelligence artificielle**. Dans ce cadre pluridisciplinaire, vous utiliserez aussi les **domaines de l'analyse de données, des statistiques, de l'optimisation, du Machine Learning et du Deep Learning**.

Vos missions principales seront :

- d'étudier quelques algorithmes pertinents sur la prise de décision appliquée à la classification de séries temporelles. En particulier, les approches récentes de Reinforcement Learning pourront servir de base de travail ;
- d'implémenter un sous-ensemble de ces algorithmes, en utilisant autant que possible les bibliothèques proposées par la communauté (Keras, TensorFlow, Scikit-learn, etc.) ;
- de tester ces implémentations sur le jeu de données associé à l'étude;
- d'analyser les résultats obtenus et d'émettre un avis critique sur le potentiel des méthodes testées.

Profil

De formation Bac +5, vous êtes actuellement en Master ou en école d'ingénieur, avec une spécialisation en apprentissage statistique, analyse de données, traitement du signal ou en mathématiques appliquées.

Vous avez une expérience en analyse de données (projet, stage, etc.), avec **implémentation sous Python** (bibliothèques Keras, Scikit-learn etc.) d'outils de visualisation, d'algorithmes de Machine Learning et de réseaux de neurones.

Vous êtes capable de programmer efficacement, tout en veillant à ce que votre code soit évolutif et facilement réutilisable. Des connaissances en parallélisation de code seraient également appréciées.

Vous êtes autonome, rigoureux, et faites preuve d'une curiosité technique et scientifique.

Vous avez le sens de la communication écrite et orale dans un environnement multidisciplinaire.

Vous parlez couramment anglais et êtes en mesure de réaliser une étude bibliographique dans cette langue.

Entreprise

L'innovation scientifique et technologique est au cœur de la stratégie de bioMérieux. Acteur mondial dans le domaine du diagnostic in vitro depuis plus de 50 ans, et leader mondial des marchés de la microbiologie clinique et industrielle, bioMérieux conçoit, développe, produit et commercialise des solutions de diagnostic (réactifs, instruments et logiciels) destinées à des applications médicales et industrielles, qui déterminent l'origine d'une maladie ou d'une contamination pour améliorer la santé des patients et assurer la sécurité des consommateurs.

Fidèle à sa mission d'améliorer la santé publique dans le monde, bioMérieux est présente dans plus de 150 pays au travers de 41 filiales. Son siège social se situe à Marcy l'Etoile, dans la région de Lyon, en France.

Nous avons la volonté d'offrir à nos collaborateurs un environnement de travail optimal qui encourage l'esprit d'équipe, avec priorité donnée à la formation et des opportunités de développement de carrière à l'international.

Pour postuler

Envoi CV + lettre de motivation à : coralie.martinez@biomerieux.com

Lieu du stage : Marcy l'Etoile, proche Lyon, France.

Dates du stage : durée de 4 à 6 mois à partir de février 2018.

