



Offre n°2024-07512

Ingénieur - développement d'un outil d'IA pour recherche de cibles thérapeutiques dans le métabolisme

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Fonction : Ingénieur scientifique contractuel

Corps d'accueil : Ingénieur d'Etudes / Ingénieur de Recherche (IE/IR)

A propos du centre ou de la direction fonctionnelle

Le centre Inria de Lyon est le 9ème centre de recherche Inria. Créé en janvier 2022, il regroupe environ 300 personnes au sein de 17 équipes de recherche et des services supports à la recherche.

Ses équipes sont localisées à Villeurbanne, à Lyon Gerland, ainsi qu'à Saint-Etienne.

Le centre de Lyon est présent dans les domaines du logiciel, du calcul distribué et haute performance, des systèmes embarqués, du calcul quantique et de respect de la vie privée dans le monde numérique, mais aussi de la santé et de la biologie numériques.

Contexte et atouts du poste

Ce projet se situe dans le cadre de l'équipe-projet Inria AlstroSight, jointe avec l'entreprise de biotechnologie Theranexus et concerne la recherche de cibles dans les maladies génétiques du métabolisme. Il s'agit de développer un outil numérique permettant d'assembler des graphes métaboliques à partir de données disponibles dans les bases de données et les utiliser pour de l'apprentissage visant à la recherche de cibles thérapeutiques dans le métabolisme:

1) construction d'un graphe métabolique d'intérêt à partir de pathways récoltés dans les bases ouvertes de réactions métaboliques humaines par technologies du web sémantique.

2) augmentation des graphes métaboliques obtenus en les décorant par des données expérimentales spécifiques à la maladie ciblée, quel que soit leur type (transcriptomiques, protéomiques, métabolomiques...).

3) apprentissage par Deep Learning (Graph Neural Networks - GNN) pour différentes tâches, de la prédiction de liens à la classification de graphes.

4) approche d'explicabilité (Lime/Shap) pour repérer les nœuds (ou sous-parties) du graphe qui ont le plus de poids dans la classification, i.e., les cibles recherchées.

Deux cibles applicatives initiales sont prévues (données déjà disponibles): recherche de cibles pour des maladies génétiques du métabolisme avec Theranexus (données de RNASeq sur culture cellulaire) et explicabilité d'un biomarqueur métabolique pour les phases précoces de Parkinson (RMN plasmatique sur animaux et patients humains) avec le GIN (Institut des Neurosciences de Grenoble).

Le verrou principal est la faible quantité d'échantillons, dans un contexte de deep learning. L'hypothèse est que le 'préconditionnement' des données par leur positionnement sur un graphe métabolique défini au préalable devrait améliorer l'efficacité. Des approches d'apprentissage auto-supervisé seront aussi envisagées.

Cet outil, intitulé MetaboGraph, écrit en Python, et avec des standards du web sémantique (RDF) pour l'interrogation des bases de données et la création des graphes de pathways métaboliques. Il se compose de deux outils distincts et indépendants. Un premier outil permet de construire des graphes métaboliques à partir des données publiquement accessibles comme celles de Reactome ou BioPAX, et d'intégrer des informations omiques dans ce graphe. Il permet aussi de regrouper le tout au sein de datasets PyTorch Geometric. Ce premier outil est ensuite utilisé comme une bibliothèque externe par le second, qui se charge de l'apprentissage par GNN à partir de données PyTorch Geometric.

Mission confiée

Parmi les étapes 1) à 4) ci-dessus, les étapes 1) à 3) sont terminées totalement ou en partie (certaines optimisations et validations restent à faire). La mission de l'ingénieur.e recruté.e consistera à terminer

l'implémentation du code de toutes les étapes, de le valider sur des bases de données transcriptomiques ouvertes (e.g., LINCS/CMAP, <https://clue.io/lincs>), puis sur deux cas d'intérêt concrets issus de nos collaborateurs.

Le travail s'effectuera sur la supervision de Jan-Michael Rye, Ingénieur de Recherche, pour les aspects développement logiciel et technologique et d'Hugues Berry, Directeur de Recherche, pour la supervision scientifique. Il inclura à la fois une activité importante de développement de code mais aussi des interactions étroites avec nos collaborateurs qui produisent les données expérimentales (transcriptomique, métabolomique).

Principales activités

- finalisation et validation des éléments déjà implémentés au moins en partie
- utilisation pour apprentissage par Graph Neural Networks sur données ouvertes de transcriptomique (e.g., LINCS)
- exploration des tâches accessibles étant données les volumes de données disponibles: classification de graphes, classification de nœuds, inférence de liens
- implémentation de l'explicabilité (LIME/SHAP)
- validation sur des données ouvertes type LINCS
- application des données de transcriptomique de maladies génétiques du métabolisme
- exploration des possibilités d'augmentation de données et de few-shot learning avec les méthodes existantes

Compétences

- Connaissance des méthodes d'analyse de données : apprentissage supervisé, notion de réseau de neurones, notions de fouille de données.
- Connaissances des frameworks d'analyse de données : Python, Pandas, Sklearn, PyTorch etc.
- Connaissance de méthodes et outils de gestion de données massives (manipulation de bases de données). La connaissance préalable des standards du web sémantique serait un plus (RDF, Owl, Sparql)

Avantages

- Restauration subventionnée
- Transports publics remboursés partiellement
- Congés: 7 semaines de congés annuels + 10 jours de RTT (base temps plein) + possibilité d'autorisations d'absence exceptionnelle (ex : enfants malades, déménagement)
- Possibilité de télétravail (après 6 mois d'ancienneté) et aménagement du temps de travail
- Équipements professionnels à disposition (visioconférence, prêts de matériels informatiques, etc.)
- Prestations sociales, culturelles et sportives (Association de gestion des œuvres sociales d'Inria)
- Accès à la formation professionnelle
- Sécurité sociale

Informations générales

- **Thème/Domaine** : Neurosciences et médecine numériques
Instrumentation et expérimentation (BAP C)
- **Ville** : Villeurbanne
- **Centre Inria** : [Centre Inria de Lyon](#)
- **Date de prise de fonction souhaitée** : 2024-06-03
- **Durée de contrat** : 12 mois
- **Date limite pour postuler** : 2024-05-01

Contacts

- **Équipe Inria** : [AISTROSIGHT](#)
- **Recruteur** :
Berry Hugues / hugues.berry@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

L'essentiel pour réussir

Intérêt fort pour les applications à la biologie/médecine

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

Les candidatures doivent être déposées en ligne sur le site Inria.fr
Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.