



Registre des entreprises, projet sur l'apprentissage automatisé et recherche opérationnelle

Description

L'équipe du Registre des entreprises (RE) de la Division des méthodes d'enquête auprès des entreprises est à la recherche d'un étudiant pour effectuer des travaux de recherche et développement afin d'améliorer la classification industrielle des entreprises canadiennes. Le RE de Statistique Canada étudie actuellement de nombreuses possibilités relatives aux mégadonnées, plus précisément aux données administratives. Ces données peuvent comprendre des descriptions relatives à l'activité industrielle qui sont vagues, incomplètes ou qui peuvent varier au fil du temps. Afin d'utiliser les données administratives à des fins statistiques, les descriptions doivent être converties en codes industriels normalisés, ce qui représente un défi en raison du volume considérable de données, de la disponibilité et de la qualité de l'information auxiliaire.

L'étudiant devra faire preuve d'innovation afin d'élaborer un processus de classification de l'activité industrielle des entreprises selon le Système de classification des industries de l'Amérique du Nord (SCIAN) actuel. Plus particulièrement, l'étudiant devra mettre à l'essai différentes techniques d'apprentissage automatisé pour classer les données de description industrielle, évaluer les résultats de ces techniques, déterminer une ou plusieurs techniques (le cas échéant) menant à des résultats d'une qualité suffisante afin d'être utilisées pour la production régulière. Le poste offre une occasion unique de travailler sur un projet d'une grande visibilité. On s'attend à ce que l'étudiant contribue de manière importante à l'élaboration, à l'application et à l'évaluation des techniques d'apprentissage automatisé dans un élément clé du RE. Il est donc essentiel de maîtriser les connaissances appropriées en la matière. On demandera à l'étudiant de documenter les techniques utilisées et les résultats obtenus. Le travail sera effectué au sein d'une équipe multidisciplinaire comprenant des statisticiens, des économistes et des programmeurs.

Si le temps le permet, on pourrait demander à l'étudiant d'étudier la possibilité d'intégrer l'apprentissage automatisé aux processus de couplage d'enregistrements du RE.

Connaissances requises

- Compétences avancées en langages de programmation pour effectuer de l'apprentissage automatisé comme R, Python ou JAVA, ainsi qu'en langages de programmation statistique comme SAS et SAS Enterprise Miner;
- Connaissances des techniques d'apprentissage automatisé.

Durée

Quatre (4) mois, avec possibilité de renouvellement. Le mandat commencera au début de 2018, selon la disponibilité de l'étudiant.

Pour postuler :

Les personnes intéressées à poser leur candidature doivent envoyer leur CV à jour, y compris les coordonnées d'une personne de référence, à Javier Oyarzun (javier.oyarzun@canada.ca), méthodologiste principal à Statistique Canada.