

STAGE

Analyse des données télé-billettiques à l'aide de modèles d'apprentissage pour la caractérisation du fonctionnement d'un système de transport multimodal

CONTEXTE de l'IRT SYSTEMX

Au sein de SystemX, Institut de Recherche Technologique dans le domaine de l'ingénierie numérique des systèmes, regroupant sur ses projets de recherche près de 300 collaborateurs, vous serez partie prenante du projet **IVA** (Information Voyageur Augmentée) qui vise à analyser et modéliser les comportements des voyageurs en situation nominale et perturbée et à mesurer l'impact de l'information voyageur sur les schémas de mobilité. Votre stage s'inscrit dans un cadre d'une activité de recherche commune à plusieurs partenaires du projet IVA : Ile-de-France-Mobilités (ex STIF), SNCF, IFSTTAR, Kisio digital et SpirOps.

Ce stage aura pour principal objectif d'**étudier l'intérêt d'utiliser des modèles de prédiction individuels de déplacement** (chaîne de Markov [1], modèle Bayésien naïf régularisé, modèle à base de factorisation de tenseur [2]) pour effectuer dans un second temps des **détections d'anomalies au sein d'un réseau de transport**. Différents modèles de prédiction devront être mis en œuvre et comparés. Le stage s'attachera également à **développer et comparer plusieurs stratégies** de détection à base de tests statistiques, par exemple. Une fois cette étape de détection effectuée, l'**analyse des comportements collectifs** de changement d'itinéraires des usagers impactés (filtrage de la sous population impactée) par l'anomalie devra être menée.

[1] Unraveling traveler mobility patterns and predicting user behavior in the Shenzhen metro system, Chao Yang, Fenfan Yan & Satish V. Ukkusuri, *Transportmetrica A: Transport Science* 2017.

[2] Anomaly detection in smart card logs and distant evaluation with Twitter : a robust framework, Emeric Tonnelier, Nicolas Baskiotis, Vincent Guigue, and Patrick Gallinar, *Neurocomputing* 2018.

<https://www.irt-systemx.fr/project/iva/>

VOS MISSIONS

- Etat de l'art méthodologique et applicatif
- Développement de plusieurs modèles de prédiction individuels et comparaisons de leurs performances
- Mise en œuvre et évaluation de stratégies de détection d'anomalie au sein d'un réseau de transport
- Analyse des comportements collectifs de changement d'itinéraires d'usagers réguliers impactés par une perturbation du réseau de transport

PROFIL RECHERCHÉ

Etudiant(e) en fin de cycle Ingénieur ou Master2 dans le domaine de la science des données, du *machine learning*, statistiques, mathématiques appliquées ou informatique, avec une spécialisation dans la science des données / machine learning

VOS COMPETENCES & APPTITUDES

- Bon niveau de maîtrise dans l'un des langages de programmation (Python, R)
- Machine Learning, statistiques
- Base de données
- Esprit d'analyse, autonomie, travail collaboratif, et intérêt pour la recherche appliquée
- Aptitude à communiquer aussi bien à l'oral qu'à l'écrit (en français et anglais)

INFORMATIONS

- Durée : 6 mois
- Stage rémunéré
- Participation aux frais de restauration à la cantine entreprise, et aux frais de transport en commun