

Sujet de stage École ingénieur / M2

CIRAD - UMR TETIS, Montpellier, France

Titre

Apprentissage automatique et profond pour la prédiction d'indicateurs de sécurité alimentaire à partir de données hétérogènes

Keywords

Apprentissage Automatique, Apprentissage Profond, Sécurité alimentaire, Données Hétérogènes

Context

La sécurité alimentaire (SA) est un problème central dans de nombreuses régions du monde, comme en témoigne sa présence parmi les 17 objectifs de développement durable (ODD 2 - Faim zéro). Pour surveiller les situations d'insécurité alimentaire, plusieurs systèmes d'alerte précoce sont actifs aujourd'hui, tels que GIEWS (FAO) et FEWSNET (USAID). Ces systèmes utilisent un ensemble limité de types de données, i.e., données agroclimatiques provenant d'images satellites et indicateurs extraits d'enquêtes auprès des ménages sur les facteurs nutritionnels, économiques et liés à la production. De plus, une intervention humaine est souvent nécessaire pour combiner et résumer toutes les sources d'information.

Les recherches menées précédemment par TETIS ont montré que des données ouvertes hétérogènes, liées à différents niveaux à la sécurité alimentaire, peuvent être utilisées pour développer des méthodes d'apprentissage automatique capables de produire des indicateurs en lien direct avec la SA. L'idée est d'utiliser des données qui représentent des proxys pour les raisons multiples et interdépendantes à l'origine de ce phénomène. Quelques exemples peuvent être des rasters d'informations spatiales (i.e. cartes de distance aux infrastructures, aux marchés, etc.), information géographique bénévole (Open Street Map), données météorologiques (i.e. températures moyennes, cumuls de précipitations, etc.), indicateurs économiques et données textuelles issus des médias locaux [2,3].

L'objectif de ce stage est d'évaluer le framework développé à TETIS [1] et testé initialement sur le cas du Burkina Faso, sur différents pays d'Afrique, notamment Rwanda et Tanzanie. L'idée à la base du framework est d'exploiter des stratégies d'apprentissage automatique et profond de dernière génération qui peuvent faire face à des aspects critiques, tels qu'une grande hétérogénéité des données d'apprentissage, mis face à une quantité des données de vérité de terrain (scores de SA) limitée.

Dans un premier temps, le travail sera focalisé sur l'actualisation des sources d'un ensemble de données hétérogènes adaptées aux zones d'étude. Cela inclut la collecte et la mise en lien des données à différentes échelles et de nature différentes.

Une fois finalisée la base des données, l'étudiant s'occupera de tester le framework d'entraînement des modèles et l'évaluation des modèles d'estimation des indices de SA ainsi entraînés. Les résultats attendus sont soit quantitatifs (évaluation des performances du framework par rapport à des données de référence), soit qualitatifs (production des cartes d'insécurité alimentaire à l'échelle nationale).

Compétences du candidat/e :

- Analyse des données (collecte, exploration, mise en lien)

- Programmation (préférentiellement Python)
- Capacité d'analyses, rédactionnelles et de synthèse
- Travail d'équipe

Informations complémentaires :

Durée de 6 mois, à partir de février 2024.

Le stage se déroulera au CIRAD, dans l'UMR TETIS (Territoire, Environnement, Télédétection et Information Spatiale), située dans les locaux de la Maison de la Télédétection à Montpellier.

Le stagiaire sera encadré par Simon Madec, Jeremy Lavarenne et Roberto Interdonato (CIRAD, UMR TETIS).

Financement:

Le stage est soutenu par le projet SCOSSA (Science des données hétérogènes pour la Sécurité Alimentaire) financé par le Centre national d'études spatiales (CNES). La rémunération du stagiaire sera de l'ordre de 600 euros par mois.

Modalité de candidature :

Attention : cette proposition ne concerne que les stages d'étudiants sous convention avec une école ou une université : il ne s'agit pas d'une offre d'emploi.

Envoyer une lettre de motivation d'une page, précisant en outre la durée et période possible du stage, un CV détaillé, et un relevé des notes à : simon.madec@cirad.fr , jeremy.lavarenne@cirad.fr , roberto.interdonato@cirad.fr

Date limite pour l'envoi du dossier : 30 Novembre, 2023

Bibliographie

[1] Hugo Deléglise, Roberto Interdonato, Agnès Bégué, Elodie Maître d'Hôtel, Maguelonne Teisseire, Mathieu Roche. Food security prediction from heterogeneous data combining machine and deep learning methods. Expert Syst. Appl. 190: 116189 (2022)

[2] Cheick Tidiane Ba, Chloé Choquet, Roberto Interdonato, Mathieu Roche. Explaining food security warning signals with YouTube transcriptions and local news articles. GoodIT 2022: 315-322

[3] Hugo Deléglise, Agnès Bégué, Roberto Interdonato, Elodie Maître d'Hôtel, Mathieu Roche, Maguelonne Teisseire. Mining News Articles Dealing with Food Security. ISMIS 2022: 63-73