

## Stage

**Intitulé : Fusion des données multi-capteurs**

**Lieu du stage:**

Equipe DIAMS, LIS, Campus St Jérôme, Aix-Marseille Université Avenue Normandie-Niemen, Marseille

**Descriptifs et objectifs :**

La fusion de données est un ensemble de méthodes visant à créer des indicateurs en agrégeant des données provenant de sources hétérogènes. Les domaines d'application sont multiples tels que le domaine médical [1], la robotique, la localisation [2] et les smart cities [3].

En effet, l'utilisation de la combinaison des observations d'un même phénomène peut améliorer l'estimation de la valeur observée. Plusieurs techniques d'apprentissage automatique peuvent être utilisées telles que les machines à vecteur de support, les méthodes ensemblistes (e.g. les forêts aléatoires) ou encore l'apprentissage profond.

Nous nous plaçons dans le domaine du climat (Smart Environment) où les données sont multi-sources, hétérogènes et multimodales incluant des capteurs, et des référentiels (e.g., zones inondables) et nous nous intéressons à la fusion orientée prédiction qui combine les sorties de plusieurs classifieurs.

L'objectif de ce stage est donc de proposer une ou plusieurs méthodes pour la fusion des prédictions de plusieurs classifieurs afin de prévenir des inondations dans les zones potentiellement à risque.

**Résultats attendus:**

Pour cela l'étudiant devra :

- Réaliser un état de l'art sur les techniques de fusion des prédictions [3-7]
- Réaliser une étude expérimentale des différents modèles existants avec des jeux de données réels (en particulier <https://iee-dataport.org/keywords/multisensor-data-fusion>)
- Proposer un modèle et un cadre pour la fusion des prédictions de données multi-sources
- Mettre en place l'architecture et les méthodes proposées

**Prérequis:**

- Connaissances en sciences des données, statistiques et apprentissage
- Solides compétences en développement en Python.

**Encadrants :** Etienne Thuillier, Laure Berti-Equille et Sana Sellami

**Adresse(s) courriel(s) :** [etienne.thuillier@lis-lab.fr](mailto:etienne.thuillier@lis-lab.fr), [sana.sellami@lis-lab.fr](mailto:sana.sellami@lis-lab.fr), [laure.berth@ird.fr](mailto:laure.berth@ird.fr)

**Laboratoire / Équipe :** LIS/DIAMS

**Durée du stage** : 5 mois

**Localisation** : St Jérôme

**Gratification** : Environ 500 euros/mois (<https://www.service-public.fr/simulateur/calcul/gratification-stagiaire> )

**Références :**

- [1] Muzammal, Muhammad & Talat, Romana & Sodhro, Ali & Pirbhulal, Sandeep. (2020). A Multi-sensor Data Fusion Enabled Ensemble Approach for Medical Data from Body Sensor Networks. Information Fusion. 53. 155-164. 10.1016/j.inffus.2019.06.021.
- [2] Valentin Magnier, Fusion de données multi-capteurs pour l'estimation de la zone navigable pour le véhicule à conduite automatisée, Thèse Univ. Paris Saclay 2018. <http://www.theses.fr/2018SACLE003>
- [3] Billy Pik Lik Lau, Marakkalage Sumudu Hasala, Yuren Zhou, Naveed Ul Hassan, Chau Yuen, Meng Zhang, U-Xuan Tan: A survey of data fusion in smart city applications. Inf. Fusion 52: 357-374 (2019)
- [4] Fei Yang, Yixin Hua, Xiang Li, Zhenkai Yang, Xinkai Yu, Teng Fei, A survey on multisource heterogeneous urban sensor access and data management technologies, Measurement: Sensors, (2021)
- [5] P. Benedetti, R. Gaetano, K. Ose, R. Pensa, S. Dupuy, D. Ienco. M3Fusion : Un modèle d'apprentissage profond pour la fusion de données satellitaires Multi-Echelles/Modalités/Temporelles ([https://agritrop.cirad.fr/589152/1/CFPT2018\\_paper\\_benedetti-1.pdf](https://agritrop.cirad.fr/589152/1/CFPT2018_paper_benedetti-1.pdf))
- [6] Tong Meng, Xuyang Jing, Zheng Yan, Witold Pedrycz: A survey on machine learning for data fusion. Inf. Fusion 57: 115-129 (2020)
- [7] J. Zhao , X. Xie , X. Xu , S. Sun , Multi-view learning overview: recent progress and new challenges, Inf. Fusion 38 (2017) 43–54