

APPRENTISSAGES FEDERES SUR LES DONNEES DE MOBILITE

POST-DOC - 2 ANS A TOULOUSE

Contexte et problématique

Le projet VILAGIL porte sur de multiples directions permettant de créer un écosystème capable de répondre aux besoins de la mobilité pour la région Occitanie. Dans le cadre de ce projet, l'action « Data et Mobilité » vise à développer des mécanismes d'intégration automatique, de stockages multi-stores et d'accès fédérés aux données. Ces données seront fournies par les partenaires non académiques du projet tels que Toulouse Metropole, Sicoval, Tisséo Collectivité. Dans certains cas ces données seront distribuées sur les différents sites des partenaires sans possibilité des les intégrer de manière centralisée. L'offre de postdoc cible donc particulièrement cette problématique en proposant des approches d'apprentissages fédérés afin de permettre l'exploitation de toutes ces données.

Objectifs

L'objectif de ce poste est de développer des mécanismes d'agrégation dans un contexte d'apprentissage fédéré appelé aussi « Federated Learning (FL)» [1][2]. En FL, les organismes peuvent participer aux tâches d'apprentissage sans forcément partager leurs données, il leur suffit de partager les paramètres locaux des modèles d'apprentissage. Il est néanmoins pas évident à ce jour de fournir des mécanismes d'agrégation en mode fédéré qui prouvent leurs efficacités en présence de données hétérogènes. Dans la littérature, certaines méthodes d'agrégation sont proposées telles que FedAvg [3], FedMA [4] ou FedPer [5]. Ces méthodes sont limitées et ne permettent pas de prendre en compte de manière efficace des masses de données fortement hétérogènes. Nous souhaiterons, vu la forte hétérogénéité des données qui seront collectées dans VILAGIL (données de capteurs, images, vidéos, etc.), pouvoir proposer des mécanismes qui répondent efficacement à différents scénarios d'agrégation pour les différents types d'apprentissages.

Références Bibliographiques

[1] Qiang Yang, Yang Liu, Tianjian Chen, Yongxin Tong. Federated Machine Learning: Concept and Applications. ACM Transactions on Intelligent Systems and Technology 10(2):1-19. 2019.

[2] Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A.N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., D'Oliveira, R.G., Rouayheb, S.E., Evans, D., Gardner, J., Garrett, Z.A., Gascón, A., Ghazi, B., Gibbons, P.B., Gruteser, M., Harchaoui, Z., He, C., He, L., Huo, Z., Hutchinson, B., Hsu, J., Jaggi, M., Javidi, T., Joshi, G., Khodak, M., Konečný, J., Korolova, A., Koushanfar, F., Koyejo, O., Lepoint, T., Liu, Y., Mittal, P., Mohri, M., Nock, R., Özgür, A., Pagh, R., Raykova, M., Qi, H.,

Ramage, D., Raskar, R., Song, D.X., Song, W., Stich, S.U., Sun, Z., Suresh, A.T., Tramèr, F., Vepakomma, P., Wang, J., Xiong, L., Xu, Z., Yang, Q., Yu, F.X., Yu, H., & Zhao, S. (2019). Advances and Open Problems in Federated Learning. ArXiv, abs/1912.04977.

[3] MCMAHAN, Brendan, MOORE, Eider, RAMAGE, Daniel, et al. Communication-efficient learning of deep networks from decentralized data. In : Artificial Intelligence and Statistics. PMLR, 2017. p. 1273-1282

[4] ARIVAZHAGAN, Manoj Ghuhan, AGGARWAL, Vinay, SINGH, Aaditya Kumar, et al. Federated learning with personalization layers. arXiv preprint arXiv:1912.00818, 2019.

[5] ARIVAZHAGAN, Manoj Ghuhan, AGGARWAL, Vinay, SINGH, Aaditya Kumar, et al. Federated learning with personalization layers. arXiv preprint arXiv:1912.00818, 2019.

Encadrement et environnement

- **Laboratoire** : Institut de Recherche en Information de Toulouse, IRIT, CNRS/UMR5505
- **Direction scientifique** :
Imen Megdiche, Equipe SIG (Imen.Megdiche@irit.fr)
André Péninou, Equipe SIG (Andre.Peninou@irit.fr)
Olivier Teste, Equipe SIG (Olivier.Teste@irit.fr)
- **Lieu de travail** : IRIT Site Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse, FRANCE
- **Durée** : 2 ans
- **Démarrage** : possible dès Mars 2021

Dossier de candidature

Nous recherchons un(e) candidat(e) motivé(e), force de propositions, avec un solide background en mathématiques appliquées et des connaissances confirmées en IA.

Les candidatures sont à envoyer à (imen.megdiche@irit.fr) et (olivier.teste@irit.fr) . Le dossier doit contenir :

- CV incluant les publications scientifiques
- Lettre de motivation
- Copie du dernier diplôme
- Pré-rapport et rapport de soutenance de thèse