AMBASSADE DE FRANCE EN INDE
Liberté Égalité Fraternité

CEFIPRA

INSTITUT FRANÇAIS India

Job offer – PhD in Machine Learning

Research Project Short Title as Submitted to CEFIPRA: "Fully Asynchronous Distributed Learning with Adversaries: Theory and Application"
Principal Investigator contact: "Alexandre Reiffers-Masson, alexandre.reiffers-masson@imt-atlantique.fr", IMT Atlantique
Reference Number of the Job Offer: IFI_CEF_24_08

Project description

- **Keywords:** Vertical federated machine Learning, Learning with Adversaries, Stochastic Approximation, L1-minimisation

- **Context:**

  Training with distributed data using cluster-based large-scale learning, federated learning, and edge computing is popular in decentralized machine learning and multi-agent Reinforcement Learning (RL). This has made the design and analysis of distributed stochastic optimization algorithms---that are i.) asynchronous, ii.) robust to network faults and malicious attacks, iii.) fast, iv.) communication efficient, and v.) differentially private ---a hot research topic.
  Traditional methods do not account for failure or malicious attacks. A few recent ones do, and they can be subdivided into two groups: synchronous and asynchronous. In synchronous methods, at every iteration, the estimates (e.g., gradient) from the different worker nodes are sought simultaneously. To achieve resilience against malicious attacks, these estimates are then aggregated using robust measures. While these methods are simple and easy to implement, their performance degrades in practical settings where agents often go offline and/or compute/communicate at different rates. This brings forth the need for asynchronous distributed learning robust to the presence of adversaries.

- **Abstract of the Research Project:**

  Our recent work (Ganesh et al., 2023) proposes a novel asynchronous algorithm to solve a distributed stochastic linear equation in adversarial settings. Our proposed project is to extend this idea to distributed stochastic (possibly non-convex) optimization. More specifically we are interested in working in the context of vertical federated learning. Vertical federated learning is a specific scenario of federated learning, in which all clients own the same dataset but each client has a unique set of features. One can think of this set-up as learning with distributed features. Vertical FL is common in e-commerce, financial, and healthcare applications, which are more sensitive to privacy leakage and adversarial attacks. The goal of this Ph.D. thesis will be to develop new learning algorithms asynchronous for vertical FL with adversaries, based on our previous work (Ganesh et al., 2023), and also study their performances mathematically and empirically.
- **Scientific Objectives of the Project:**

  - Distributed Stochastic Linear Equation: Discuss variants of our algorithm in (Ganesh et al., 2023) so that we also obtain the convergence rate, and explore the idea of using momentum to accelerate the current algorithm.
  - Distributed Stochastic (Non-Convex) Optimization - Extend the above algorithm and the analysis to minimize a general (possibly non-convex) function □.
  - Design of Robust Observation Matrices and Applications: Identify efficient ways in which the distribution among entities of features can be designed from scratch so that it satisfies our robustness criterion. Similarly, design algorithms so that a given distribution among entities of features can be transformed to make it robust.

- **Methodology and Timeline of the Project:**

    1. 1st Year: Distributed Stochastic Linear Equation for Vertical Federated Learning
        a. Convergence rate and lower bounds
        b. Acceleration
        c. Communication-Efficiency and Privacy
    2. 2nd Year: Distributed Stochastic (Non-Convex) Optimization for Vertical Federated Learning
        a. Almost sure Convergence
        b. Speed, Communication-Efficiency, and Privacy
    3. 3rd Year: Design of Robust Observation of Features Distribution
        a. Efficient aggregation for robust observation matrix
        b. Expansion of the observation matrix

(Ganesh et al., 2023) Ganesh, Swetha, Alexandre Reiffers-Masson, and Gugan Thoppe. "Online Learning with Adversaries: A Differential-Inclusion Analysis." *2023 62nd IEEE Conference on Decision and Control (CDC)*. IEEE, 2023.

Candidate profile

- Only Indian candidates or candidates with a research experience in India are eligible; French candidates are not eligible
- Applicants for PhD must have a master's degree (or be in the process of obtaining one) or have a University degree equivalent to a European Master's (5-year duration) to be eligible at the time of the deadline of the call;
- No competencies in French language is required
- Candidate competences: Solid mathematical background, some experience in machine learning and possible (but not mandatory) in federated machine learning
- Candidate know-how: Strong computer science skills in Python, Good communication skills, excellent written English skills, Motivated, inquisitive and self-driven
- Expected starting date: **01-10-2024**

How to candidate?

Documents to be provided :

i. A cover letter (reasons for the candidature, professional project …) max 2 pages
ii. A copy of the master's degree or a proof of the program followed (and expected date of end) OR A copy of the PhD degree or a proof of the PhD program followed (and expected date of defense) max 1 page
iii. A copy of results for previous scholarship (max 3 pages)
iv. International curriculum vitae (max 2 pages)
v. Two letters of recommendation: one from any Indian institution and one from the French institution planned to host the candidate –mandatory- (max 2 pages)
vi. All should be submitted within 1 pdf file of no more than 10 pages.

Applications should be submitted to the following email address: msi@ifindia.in mentioning the reference number of the Job offer clearly.

Research Project Title as Submitted to CEFIPRA: "Fully Asynchronous Distributed Learning with Adversaries: Theory and Application"

Candidates are requested to contact the French scientific principal investigator of the project before submission. A recommendation letter from the scientific principal investigator is mandatory.

Benefits:

- Monthly allowance of 1710 euros for PhD
- Travel allowance
- University fee
- Carte de séjour fee
- Campus France management fee
- Registration to the French social security scheme

Selection process:

Selection is made by a dedicated selection committee of 4 personness (2 members of the Embassy of France in India and 2 external experts). Decisions will be transmitted to CEFIPRA. No consideration will be given for candidates with no recommendation letter from the French institution.

Criteria for applicants' selection:

Academic excellence

- Excellence of the Academic background, Academic records, Honors, Letters of support, Participation to international research projects, exchange programmes and conferences.

Motivation and qualities

- Academic maturity: appropriation of the thesis project (stakes and contexts) • Quality of the presentation (oral expression, skills for synthesis, English level) • Maturity of the professional project: capacity to project her/himself within five years in terms of career development.

About CEFIPRA:
Indo-French Center for the Promotion of Advanced Research (CEFIPRA/IFCPAR) is an Indian body which promotes scientific cooperation between France and India in advanced fields of Science and Technology. It is supported by the Department of Science and Technology, Government of India and the Ministry of Europe and Foreign Affairs of the French government