

Pedro Valdeira

PH.D. CANDIDATE · ELECTRICAL AND COMPUTER ENGINEERING

✉ pvaldeira@cmu.edu | 🏠 valdeira.github.io | 📧 Valdeira | 🌐 pvaldeira | 🎓 Google Scholar

RESEARCH INTERESTS

I am broadly interested in developing **efficient machine learning** methods for real-world settings, with a particular interest in **optimization** methods for **distributed machine learning**. I am currently working on **provable methods** for efficient and scalable **vertical federated learning**, where multiple entities collaboratively train a model on feature-distributed samples without sharing their local data (for example, companies holding different private attributes about shared users).

EDUCATION

Carnegie Mellon University

PH.D. CANDIDATE IN ELECTRICAL & COMPUTER ENGINEERING

- Advisor: [Yuejie Chi](#)

Pittsburgh, PA, US

Sept 2020 - present

Instituto Superior Técnico

PH.D. CANDIDATE IN ELECTRICAL & COMPUTER ENGINEERING

- Advisors: [João Xavier](#) and [Cláudia Soares](#) (co-advisor)

Lisbon, Portugal

Sept 2020 - present

Carnegie Mellon University

MS IN ELECTRICAL & COMPUTER ENGINEERING

- Selected graduate coursework: Machine Learning, Deep Learning, Optimization, Statistics

Pittsburgh, PA, US

Sept 2020 - Dec 2022

Instituto Superior Técnico

INTEGRATED MS IN AEROSPACE ENGINEERING

- Thesis: “Distributed Learning and Inference for Gaussian Mixtures” (20/20). Advisors: [João Xavier](#), [Cláudia Soares](#)
- Elected Student Rep: managed exam calendars, handled schedule conflicts, liaised between students & teachers

Lisbon, Portugal

Sept 2014 - Nov 2019

Tsinghua University

GRADUATE STUDENT IN COMPUTER SCIENCE & TECHNOLOGY (EXCHANGE PROGRAM)

Beijing, China

Feb 2018 - Jun 2018

RESEARCH & INDUSTRY EXPERIENCE

Yuejie Chi Group

PH.D. RESEARCHER

- Designed novel algorithms for efficient vertical federated learning

Pittsburgh, PA, US

Sept 2020 - present

Institute for Systems and Robotics, IST

PH.D. RESEARCHER IN THE SIGNAL & IMAGE PROCESSING GROUP

- Designed novel algorithms for efficient vertical federated learning

Lisbon, Portugal

Sept 2020 - present

GRADUATE RESEARCHER IN THE SIGNAL & IMAGE PROCESSING GROUP

- Developed approximate inference and learning methods for data distributed by features over graphs

Jan 2020 - Aug 2020

Critical Software

INTERN

- Contributed to the development of a proof-of-concept IoT data acquisition pipeline
- Implemented Arduino-to-RPi data transfer, data processing, and communication of summary metrics to a server

Lisbon, Portugal

Jul 2017 - Sept 2017

TEACHING EXPERIENCE

CMU 18-661 (Introduction to ML for Engineers)

TEACHING ASSISTANT & GUEST LECTURER

- Taught recitations, held office hours, created course material (homework and exams), graded
- Gave a lecture on federated learning and presented research on vertical federated learning

Pittsburgh, PA, US

Jan 2023 - May 2023

TEACHING ASSISTANT

- Taught recitations, held office hours, created course material (homework and exams), graded

Aug 2021 - Dec 2021

PUBLICATIONS

- **Pedro Valdeira**, Shiqiang Wang, Yuejie Chi. “Vertical Federated Learning with Missing Features During Training and Inference”, *International Conference on Learning Representations (ICLR) 2025*, accepted
 - This work addresses a longstanding challenge in vertical federated learning: performing training and inference when some clients have missing features. We propose a simple yet effective approach relying on the strategic sharing of model parameters and on task-sampling to train a family of predictors. We prove the convergence of our method and demonstrate, through experiments, an improvement in accuracy of up to 18.2% in the presence of missing features.
- **Pedro Valdeira**, João Xavier, Cláudia Soares, Yuejie Chi. “Communication-efficient Vertical Federated Learning via Compressed Error Feedback”, *IEEE Transactions on Signal Processing*, 2025, accepted
 - This work tackles a key bottleneck in federated learning (FL): communication efficiency. We propose a method for vertical FL that employs lossy compression and an error feedback mechanism. Our approach matches the convergence rate of uncompressed methods without requiring the compression error to vanish, overcoming a limitation of existing methods that necessitate gradually reducing compression. Experiments show significant performance gains over prior art.
 - A short version of this work was presented as an **invited** paper at *EUSIPCO 2024*
- **Pedro Valdeira**, Yuejie Chi, Cláudia Soares, João Xavier. “A Multi-Token Coordinate Descent Method for Semi-Decentralized Vertical Federated Learning”, *in submission*
 - Most FL works consider either client-server or fully decentralized approaches. Yet the former is prone to server bandwidth bottlenecks and the latter sees slow convergence in large or sparse networks (even failing to converge in disconnected graphs). To address this, we propose a semi-decentralized vertical FL method leveraging both client-server and client-client links. By tuning the reliance on client-server links, our method can outperform the baselines in various applications.
 - A preliminary version of this work was presented at *FL NeurIPS 2022 Workshop*
- **Zhize Li**, **Pedro Valdeira**, Yuejie Chi. “Federated Optimization: Efficiency, Resiliency, and Privacy”, *in preparation*
 - This paper reviews recent literature on optimization methods for federated learning, condensing insights from key breakthroughs and highlighting important open problems that remain unsolved.
- **Pedro Valdeira**, Cláudia Soares, João Xavier. “Decentralized EM to Learn Gaussian Mixtures from Datasets Distributed by Features”, presented at *31st European Conference on Operational Research*, 2021
 - This work introduces a novel, consensus-based adaptation of the classic expectation–maximization algorithm to estimate a mixture of Gaussians in settings where the data is distributed by features across the nodes of a communication graph.

SERVICE & AWARDS

- **Mentorship:**
 - Luiz Eduardo Leite Filho, MS in Data Science at IST (*Dec 2024 – present*)
 - Francisco Freitas, MS in Computer Science at NOVA (*Sept 2024 – present*)
- **Peer Reviewership:**
 - **Journals:** IEEE Transactions on Parallel and Distributed Systems
 - **Conferences:** AISTATS (2024), EUSIPCO (2024)
- **Awards:**
 - Best MS Thesis, Portuguese branch of the International Association for Pattern Recognition, 2nd Place (*2020*)
 - Diploma of Academic Excellence, Instituto Superior Técnico (*2019*)

SKILLS & NON-FORMAL EDUCATION

- **Software:** Python, PyTorch, NumPy, LaTeX
- **Languages:** English (fluent), Portuguese (native), Mandarin (elementary proficiency, HSK4)
- **Summer school:** Oxford Machine Learning Summer School (*Jul - Aug 2021*)