Gabriel Franco

LinkedIn : www.linkedin.com/in/gaabrielfranco
Github : https://www.github.com/gaabrielfranco
Personal Webpage: https://cs-people.bu.edu/gvfranco/
E-mail: gvfranco@bu.edu
Mobile: +1 (617) 685-9112
Address: 665 Commonwealth Ave, Boston, MA

EDUCATION

Boston University, Boston, MA
PhD, Computer Science
Sep 2021 - Present
Advisor: Mark Crovella

Federal University of Viçosa, Brazil
MSc, Computer Science
Aug 2018 - Jun 2021
Advisor: Giovanni Comarela

Federal University of Viçosa, Brazil
B.S., Computer Science
Mar 2014 - July 2018

RESEARCH INTERESTS

Mechanistic Interpretability, Large Language Models, Machine Learning, Weakly Supervised Learning

SCHOLARSHIP

Boston University Research Scholarship, PhD Student Sep 2021 - Present
Worked on evaluation of weakly supervised learned classifiers. Currently, working on mechanistic interpretability of Large Language Models (LLMs).

CAPES Research Scholarship, Master Student Aug 2018 - Jul 2020
Worked on better cross-validation strategies for a weakly supervised learning problems.

CNPq Research Scholarship (PIBIC/CNPq) Aug 2017 - Jul 2018
Worked with improving the malaria model on Autosimmune, a multiagent human immune system simulator made in JAVA.

CNPq Research Scholarship (PIBIC/CNPq) Aug 2015 - Jul 2016
Developed the Bio-Oracle software. Bio-ORACLE is a software written in JAVA which uses data mining techniques to help decision-making in bioethics. Bio-ORACLE was developed at the Laboratory of Epidemiological and Computational Methods in Health of Department of Medicine and Nursing in Federal University of Viçosa.

FAPEMIG Research Scholarship (PIBIC/FAPEMIG) Mar 2015 - Jul 2015
Worked with modeling the plasmodium on Autosimmune, a multiagent human immune system simulator made in JAVA. Autosimmune was developed at the Laboratory of Epidemiological and Computational Methods in Health of Department of Medicine and Nursing in Federal University of Viçosa.

INDUSTRY EXPERIENCE

Data Scientist Intern at Microsoft May 2024 - Aug 2024

- Fine-tuned multi-modal (text and image) Small Language Models (SLMs) to validate their feasibility for on-device execution in Windows, culminating in a technology demonstration presented by the applied science team to the CEO.

Data Scientist at SEEK Sep 2020 - Jul 2021

- Designed, implemented, and maintained recommender systems to provide personalized job ad recommendations for customers.
- Improved average response time of a recommender system by more than 50%.
- Increased business metrics for a recommender system with statistical significance after an A/B test.

Data Scientist at Localiza **Jul 2020 - Sep 2020**

- Developed machine learning models to identify possible reliable customers.
- Analyzed client behavior on the platform to show the feasibility of using our trained model for business decision-making.
- Presented results to stakeholders.

TEACHING EXPERIENCE

Boston University **Jan 2025 - May 2025**
CAS CS 132: Geometric Algorithms. Teaching Assistant.
Activities:
- Prepare and teach discussions to the students.

Boston University **Sep 2023 - Dec 2023**
CDS DS 701: Tools for Data Science. Teaching Assistant.
Activities:
- Prepare and teach discussions to the students.
- Design the practical homeworks and the midterm exam (Kaggle style competition).

Federal University of Viçosa **Nov 2019**
LaTeX short course. Instructor.

Federal University of Viçosa **Mar 2019 - Jul 2019**
INF 100 - Introduction to Programming I. Teaching Assistant.

VOLUNTEER EXPERIENCE

NoBugs: Informatics Junior Enterprise **Jan 2017 - Jan 2018**
NoBugs is a junior enterprise of the UFV Computer Science course. We made low-cost web systems to regional enterprises.

PUBLICATIONS

- Mechanistic Interpretability:
  1. Franco, Gabriel, and Mark Crovella. "Pinpointing Attention-Causal Communication in Language Models". **Accepted in NeurIPS 2025.**
  2. Franco, Gabriel, and Mark Crovella. "Sparse Attention Decomposition Applied to Circuit Tracing". https://arxiv.org/abs/2410.00340.
- LLM evaluation:
  1. Calais, Pedro, et al. "Disentangling Text and Math in Word Problems: Evidence for the Bidimensional Structure of Large Language Models' Reasoning."Findings of the Association for Computational Linguistics: ACL 2025. 2025.
- Learning from Label Proportions (LLP):
  1. Franco, Gabriel, Giovanni Comarela, and Mark Crovella. "Evaluating LLP Methods: Challenges and Approaches". https://arxiv.org/pdf/2310.

19065

2. Franco, Gabriel, Mark Crovella, and Giovanni Comarela. "Dependence and Model Selection in LLP: The Problem of Variants."Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2023.

- Others:

1. Ghaemi, Golsana, Gabriel Franco, Kazem Taram, and Renato Mancuso. "Heterogeneous Memory Benchmarking Toolkit." **Accepted in RTSS 25.**
2. Franco, Gabriel, Marcos Henrique Fonseca Ribeiro, and Giovanni Comarela. "Towards an interpretable metric for DOTA 2 players: An unsupervised learning approach."2019 8th Brazilian Conference on Intelligent Systems (BRACIS). IEEE, 2019.
3. Gomes, Andréia Patricia, et al. "Plasmodium Falciparum Infection: In Silico Preliminary Studies."Abakós 5.1 (2016): 63-83.

OTHER ACADEMIC PRODUCTIONS

1. Siqueira-Batista, Rodrigo, et al. Parasitologia: Fundamentos e Prática Clínica. Guanabara, 2020. ISBN 9788527735735. (I co-wrote Chapter 4 about the computational approach in the study of parasitic diseases)
2. Comarela, G.; Franco, G. ; Trois, C. ; Liberato, A. ; Martinello, M. ; Corrêa, J. H. ; Villaça, R. Introdução à Ciência de Dados: Uma Visão Pragmática utilizando Python, Aplicações e Oportunidades em Redes de Computadores SBRC 2019 (Short Course)

SKILLS

Python; Git; TransformerLens; Numpy; Scipy; Pandas; Matplotlib; Seaborn; Scikit-learn; Statsmodel; Pytorch; Transformers; PEFT; LoRA; Fine-tuning LLM; TransformerLens; Problem Solving; Research; Machine Learning; Data Mining; Probability; Mechanistic Interpretability; Linear Algebra; C/C++

SERVICE

- Reviewer ICLR 2026
- Reviewer Mechanistic Interpretability Workshop at NeurIPS 2025
- Reviewer NeurIPS 2025
- Reviewer ICLR 2025

AWARDS & GRANTS

- KDD'23 Student Travel Award
- KDD'22 Student Travel Award

ADDITIONAL ACTIVITIES

- ACM International Collegiate Programming Contest - Regionals - 2014, 2015, 2017
- Minas Gerais State Programming Contest - 2014, 2015, 2017

RELEVANT COURSES

Boston University:

- CS542 - Machine Learning
- CS565 - Algorithmic Data Mining
- CS537 - Randomness in Computing
- CS655 - Graduate Computer Networks
- CS511 - Formal Methods 1
- DS563 - Algorithmic Techniques for Taming Big Data
- CS505 - Introduction to Natural Language Processing

• LX690 - Metrics and Evaluation in Natural Language Processing (Audited)

Federal University of Viçosa:
• INF623 - Artificial Intelligence
• INF610 - Data Structures and Algorithms
• INF723 - Data Visualization

LANGUAGES

• Portuguese (Native or bilingual proficiency)
• English (Professional working proficiency)
• Spanish (Elementary proficiency)

REFEREES

Prof. Mark Crovella
Boston University
crovella@bu.edu

Prof. Evimaria Terzi
Boston University
evimaria@bu.edu

Prof. Giovanni Ventorim Comarela
Universidade Federal do Espírito Santo
gc@inf.ufes.br