

## **Gisele Lobo Pappa (Principal Investigator)**

**Associate Professor of Computer Science**

**Computer Science Department**

**Universidade Federal de Minas Gerais**

**Google citations profile: [https://scholar.google.com/citations?user=C\\_0ZLuYAAAAJ](https://scholar.google.com/citations?user=C_0ZLuYAAAAJ)**

**Email: [glpappa@dcc.ufmg.br](mailto:glpappa@dcc.ufmg.br)/[giselepappa@gmail.com](mailto:giselepappa@gmail.com)**

**Research interests:** evolutionary computation, machine learning, applications on healthcare.

**1. Highest Academic Qualification:** Ph.D. in Computer Science. University of Kent, UK, 2007.  
(Research Area: Data Mining and Evolutionary Algorithms)

**2. Current position:** Associate Professor of Computer Science at UFMG, Brazil, since 2009. Lead of the Laboratory of Computational Intelligence (LAIC).

## **2. Publications**

- **Research-Oriented Authored Book:** Pappa, G. L., Freitas, A. A., Automating the Design of Data Mining Algorithms: An Evolutionary Computation Approach. Springer, 2010. ISBN-13: 978-3-642-02540-2.
- **Peer-Reviewed Journal Papers:** 40
- **Peer-Reviewed Full Conference Papers (about 10 pages each):** 90
- **Citations in Google Scholar** (as of 11/05/2023): 3,130 citations, H-index 30.

## **3. Community Services**

Being an active member of the evolutionary computation community, I have served as Program Chair/Co-chair for EuroGP in 2023 and 2022, and as GECCO track chair in different opportunities, sharing both the Evolutionary Machine Learning track (2022), and the Self-\* Search (2012 and 2013). I have also served as a tutorials' chair for both GECCO (2020) and PPSN (2018), and as the GECCO workshops' chair in 2015 and 2016.

I'm also active in the artificial intelligence community in Brazil and Latin-america, being the PC of the Brazilian Conference on Artificial and Computational Intelligence (BRACIS) in 2016 and the general chair for the conference in 2023.

I'm currently a member of the executive committee of ACM SIGEVO (Special Interest Group on Evolutionary Computation), and a member of the editorial board of the Springer journal Genetic Programming and Evolvable Machines and the MIT journal on Evolutionary Computation.

## **5. Supervision of PhD and Research MSc students**

I currently supervise 2 PhD and 4 Master students. I have already supervised 7 PhD students (one of them as the co-advisor) and 25 Master students.

## **7. Administrative Positions**

I have already acted as the vice-director of the Computer Science Graduate Program at UFMG and as the director of the undergraduate course of Information Systems in the same institution.

## **Research Group**

**Overview:** I'm the leader of the Laboratory of Computational Intelligence (LAIC), which is within a major research group named Speed (System Performance Evaluation and Experimental Development). Speed gathers researchers from machine learning, data mining, evolutionary computation and distributed and parallel systems. Speed currently has over 50 students, 14 of them working directly at LAIC.

LAIC is the branch of Speed that accounts for research on computational intelligence and machine learning. Our main research focuses on bringing together genetic programming and machine learning.

**Research Interests:** Currently, our main research interests are on (i) Automated Machine Learning methods and their fitness landscapes, (ii) how to put together semantics and syntax of genetic programming solutions through new representations and (iii) explainable artificial intelligence with evolutionary computation.

Within all these areas, we develop new methodologies to solve problems and apply them to real-world datasets. We currently have two major projects undergoing that bring us the data for applying the developed methods. The first is part of a big initiative named the [Innovation Center on Artificial Intelligence for Health](#) (CI-IA Health), which is a research and innovation center created to bridge the gap between computer science and the medical sciences, and is sponsored by the Brazil government together with health insurance companies. The second is a partnership with the Public Ministry of Minas Gerais, which is the institution responsible for auditing and investigating public expenses of the state of Minas Gerais. In this context, we work to detect fraud in public procurements.

**Infrastructure:** LAIC is located in the Computer Science building at UFMG in an area of over 120m<sup>2</sup>. Our infrastructure also has computer services with modern GPUs and CPUs for dealing with large models and datasets.

## **Project description**

The proposed project concerns one of our main interests, which is explainable artificial intelligence (XAI) with evolutionary computation. The area of XAI became quite popular in the past years due to current legislation that requires decisions made by algorithms to be auditable. In this sense, understanding the results provided by machine learning models that help people making these decisions became paramount. Nevertheless, most models with high accuracy are black-box, i.e. they work based on complex internal mechanisms that are mostly difficult for humans to understand. These models include artificial neural networks, support vector machines, and ensembles of different kinds.

Genetic Programming (GP) algorithms have been systematically pointed out as methods that can produce inherently interpretable models. This level of interpretability depends on the size of the model, which is something not straightforward to measure. For a long time, GP-based trees, for

example, had their simplicity measured by the number of nodes the tree model has. This is certainly a measure of complexity but it does not translate into interpretability for humans.

Measuring interpretability is hard, as it is a subjective concept. A few objective measures have been previously proposed [1,2], but they have not been correlated with a qualitative assessment of the interpretability by a human. In this direction, the main objective of this project is to propose and/or evaluate objective measures to assess the interpretability of models produced using tree-based GP, and correlate these measures with qualitative assessments made by specialists.

In order to assess and correlate these metrics with qualitative assessments, the project will be developed in partnership with the Department of Psychiatry of UFMG within the context of CI-IA Health, already mentioned above. The models will be evaluated by both psychiatrists and psychologists, and later correlated with the objective metrics. There are two candidate datasets we can choose from: the first regards predicting the risk of suicide in children and teenagers. The second regards assessing the quality of life of individuals based on questionnaires.

Regardless of the problem chosen, the results of this project will allow us to transfer the knowledge learnt to other applications of interest. The idea is for the visiting PhD student to start working with structural complexity measures, and develop a methodology for the specialists to evaluate the models. Working alongside the psychiatry department of UFMG will allow the visiting student to gain valuable insight into the interpretability of the models with the human-in-the-loop approach. This is certainly a necessary step for developing interpretable GP models.

## ***References***

- [1] Renato Miranda Filho, Anísio M. Lacerda, and Gisele L. Pappa. 2023. Explainable Regression Via Prototypes. *ACM Trans. Evol. Learn. Optim.* 2, 4, Article 14.
- [2] Brotto Rebuli, K; Giacobini, M; Silva, S; Vanneschi, L. 2023. A Comparison of Structural Complexity Metrics for Explainable Genetic Programming. To be published in *Proc. of the Genetic and Evolutionary Computation Conference (Lisbon, Portugal) (GECCO '23)*.

## **Additional information**

The [Federal University of Minas Gerais](#) is the oldest free public higher education institution in Minas Gerais. With almost a century of life, the institution is a regional and national leader in teaching, extension, culture, scientific research and generation of patents, in several areas of knowledge. Its community brings together around 72,000 people in four different cities, being the main campi in Belo Horizonte. It offers 77 undergraduate courses, 80 graduate programs, and more than 750 research centers.

The [Graduate Program in Computer Science](#) of the Federal University of Minas Gerais (UFMG) is the most well-established in the area of Computer Science in the country, and was evaluated by CAPES (the Brazilian Federal Agency for Support and Evaluation of Graduate Education) at level 7, which is the highest. Created in 1974, the program has always sought academic excellence both

in human resources and in research, publishing the results in international vehicles with high visibility.

In addition to its academic opportunities, [Belo Horizonte](#) is known for its culinary, thriving art and music scene. There are numerous museums, theatres, and galleries throughout the city, offering a rich cultural experience for visitors.

Initially, no additional funds or accommodation will be offered to the visiting student, but it is possible to apply for university accommodation. However, Brazil is not an expensive country and the SPECIES Scholarship is enough to pay the living expenses.