MARIE SKŁODOWSKA-CURIE POSTDOCTORAL FELLOWSHIPS 2025

EXPRESSION OF INTEREST FOR HOSTING MARIE CURIE FELLOWS

HOST INSTITUTION

NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Lisbon, Portugal

RESEARCH GROUP AND URL

Data Science Research Stream: https://magic.novaims.unl.pt/en/about-us/research-lines/
Data Analytics Lab: https://magic.novaims.unl.pt/en/nova-analytics-labs/nova-data-analytics-lab/

SUPERVISOR (NAME AND E-MAIL)

Leonardo Vanneschi, lvanneschi@novaims.unl.pt

SHORT CV OF THE SUPERVISOR

Leonardo Vanneschi a Full Professor at NOVA IMS, Universidade Nova de Lisboa (Portugal), and he is the Director of the Information Management research center (MagIC). With a PhD in Computer Science from the University of Lausanne (Switzerland) and a Master's degree from the University of Pisa (Italy), his academic journey includes roles at the University of Milano-Bicocca (Italy) and NOVA IMS, where he progressed from Assistant to Full Professor, and obtained the title of “Agregação” in Information Management in 2016.

Throughout his career, he has cultivated a robust international collaborative network, foundational for numerous national and international research projects. He led 4 projects and contributed to 11 more as a work package leader or researcher. The impact of his work is also reflected in almost 300 published contributions, including more than 100 in high-impact scientific journals, and numerous citations. His research interests include **Machine Learning, Data Science, Complex Systems**, and **Evolutionary Computation** and his work revolves around the vision that theoretical studies can enhance the state-of-the-art computational methods, providing practical value in real-world impactful application areas, such as **medicine, biology, engineering, economics, logistics** and many others.

Since 2011, he has held various institutional positions at NOVA IMS, including the Direction of the Data Analytics lab, the leadership of the Data Science research stream of the MagIC, the Direction of the Master course in Data Science and Advanced Analytics and the coordination of the Erasmus program. Outside NOVA IMS, he co-founded and serve on the Executive Board of the Species society (Society for the Promotion of Evolutionary Computation in Europe and its Surroundings), holds editorial roles in multiple scientific journals, and contribute to international conferences as an associated editor, steering committee member, and program committee participant. In 2015, he was honoured with the Award for Outstanding Contributions to Evolutionary Computation in Europe, in the context of EvoStar, the leading European Event on Bio-Inspired Computation. From 2018 to nowadays, he has been included every year in the list of top 2% world scientists by a study of the University of Stanford, both in the current year and careerwise.

5 SELECTED PUBLICATIONS

- L. Vanneschi. SLIM_GSGP: The Non-bloating Geometric Semantic Genetic Programming. In: Giacobini, M., Xue, B., Manzoni, L. (eds) Genetic Programming. EuroGP 2024. Lecture Notes in Computer Science, vol 14631. Springer, Cham. 2024.

- N. Rodrigues, K. M. Malan, G. Ochoa, L. Vanneschi and S. Silva. Fitness landscape analysis of convolutional neural network architectures for image classification. Information Sciences, Volume 609, Pages 711-726. 2022.

- I. Bakurov, M. Castelli, O. Gau, F. Fontanella and L. Vanneschi. Genetic Programming for Stacked Generalization. Swarm and Evolutionary Computation, Volume 65, 2021.

- L. Vanneschi and M. Castelli. Soft Target and Functional Complexity Reduction: A Hybrid Regularization Method for Genetic Programming. Expert Systems With Applications, Volume 177, 2021.
- L. Vanneschi. An Introduction to Geometric Semantic Genetic Programming. In Numerical and Evolutionary Optimization (NEO 2015), O. Schütze et al. editors. Studies in Computational Intelligence, Volume 663, pages 3–42. Springer, Cham. 2017.

PROJECT TITLE AND SHORT DESCRIPTION

Semantic and Interpretable AI

In recent years, the advent of semantic-aware Artificial Intelligence (AI) methods has marked a significant milestone. This includes some of the most advanced versions of Genetic Programming (GP), such as Geometric Semantic GP (GSGP) and its variants that leverage geometric properties of the semantic space, and the so-called Semantic Machine, the first version of a semantic-aware Artificial Neural Network. These techniques owe their success to their ability to induce a unimodal error surface, making them highly effective in optimizing training data. Combined with their notable ability to limit overfitting and promote the generation of robust models with good generalization capabilities, these techniques are highly competitive compared to many existing machine learning algorithms. This competitiveness, evidenced by numerous successful applications across various fields in recent years, motivates researchers to further extend these techniques with a new attribute: interpretability.

In a society where AI is becoming increasingly pervasive and adopted, model interpretability is becoming an essential requirement, and in some specific cases, even legally mandated. For instance, in medical applications, having models that humans can understand and interpret is crucial for their use in daily clinical practices, as well as in many other areas. The primary objective of this project is to develop new techniques that, while maintaining the characteristic of being semantic-aware with all its positive consequences, including the induction of a unimodal error surface, also have the ability to generate compact, readable models with high human interpretability. This can be achieved through the introduction of new neighborhood operators that reduce model size while maintaining competitive performance. A seminal publication presenting this type of operator appeared in 2024. However, this publication also highlighted the absolute necessity of dedicating much more research to this topic, aimed not only at better understanding the functioning of these operators and improving their usage but also at introducing new and more powerful operators that further enhance interpretability.

An integral part of this project, besides developing these innovative AI techniques, is applying these techniques to high-impact real-world problems, comparing these methods with the state-of-the-art. Among these application areas, significant developments are expected in the medical field, aimed at improving diagnostic and prognostic capabilities for various diseases, such as different types of cancer and some of the most common neurodegenerative diseases, through robust, accurate, and interpretable predictive models. Other areas of interest will certainly include radiomics and other fields focused on the analysis of features derived from medical images, as well as drug discovery, where predictive models of pharmacokinetic parameters play a fundamental role.

SCIENTIFIC AREA WHERE THE PROJECT FITS BEST*

Information Science and Engineering (ENG)