1. Short CV of the advisor

Leonardo Vanneschi

Leonardo Vanneschi is a Full Professor (Professor Catedrático) at the NOVA Information Management School (NOVA IMS), Universidade Nova de Lisboa, Portugal. His main research interests involve Machine Learning, Data Science, Complex Systems, and in particular Evolutionary Computation. His work can be broadly partitioned into theoretical studies on the foundations of Evolutionary Computation, and applicative work. The former covers the study of the principles of functioning and dynamics of Evolutionary Algorithms, with the objective of developing strategies able to outperform the state-of-the-art techniques. The latter covers several different fields among which computational biology, image processing, personalized medicine, engineering, economics and logistics.

He has published more than 300 contributions, 117 of which in internationally renown scientific journals with impact factor and 11 of which have been honoured with international awards (including 5 best paper awards at EvoStar events). He has led several research projects in the area, participating in a large number of other projects, and he is now the leader of a Laboratory on Data Analytics, the Director of the Research Center in Information Management (MagIC) at the Universidade Nova de Lisboa and the Associate Dean for Research of NOVA IMS.

His work has been consistently recognized and appreciated by the international community from 2000 to nowadays. 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. Since 2019, he has constantly been ranked by the Stanford University in the top 2% world scientists both in the current year and careerwise.

He has been keynote invited speaker in numerous national and international events. He is a member of the editorial board of three international scientific journals and of the steering committee of several international conferences, including EuroGP. He is a member of the steering board of the Species Society. He has regularly organized national and international conferences and workshops. The last one of these events was GECCO 2023, where he served as the local chair.

Since 1999, he has consistently been teaching a large number of disciplines, broadly covering the entire area of Computer Science, for several different types of students (beginners, experts, computer scientists and coming from other areas). In the last decade, his teaching activity was particularly specialized in the areas of Intelligent Systems, Machine Learning and Data Science. He has received several awards and distinctions for his pedagogical activity.

2. Description of the research group

MagIC Research Center, Nova Information Management School

The Information Management Center (MagIC) is the research unit of the Nova Information Management School (NOVA IMS), one of the world's leading institutions in the field of Information Management and Data Science. NOVA IMS consistently ranks among the top institutions globally, being placed in the TOP 5 of the Best Master's and Postgraduate Programs worldwide by Eduniversal, and it is also the first European institution to have obtained the prestigious ABET accreditation. MagIC has recently been rated as Excellent (the maximum possible evaluation) by the international evaluation panel of the Foundation for Science and Technology of Portugal (FCT). These recognitions reflect the school's strong commitment to academic excellence, innovation, and international impact.

MagIC is a dynamic and highly productive research environment, currently comprising more than 50 integrated researchers and a broad network of collaborators. Since 2011, its members have produced over 1,000 scientific publications, including more than 750 articles indexed in Web of Science and Scopus. The center's primary mission is to advance the state of the art in Information Management and Data Science through high-quality, impactful research that bridges theory and practice. MagIC is structured around two main research streams, Data Science and Management, both of which contribute to addressing complex challenges in modern data-driven environments.

Within this framework, the Data Science research stream plays a central role and is organized into four main branches: Evolutionary Computation, Deep Learning, Synthetic Data Generation, and Geoinformatics. The Evolutionary Computation branch, in which the visiting student would be integrated, focuses on both the theoretical foundations and practical applications of evolutionary algorithms. This includes the design of novel methods, the study of their properties, and their application to real-world problems in diverse domains such as healthcare, finance, and environmental systems. The group has a strong international reputation and is particularly active in areas such as semantic genetic programming, interpretable machine learning, and hybrid approaches that combine evolutionary computation with modern artificial intelligence techniques.

In 2020, NOVA IMS and MagIC further strengthened their research infrastructure by launching the Data Analytics Lab, directed by Professor Leonardo Vanneschi. This lab provides a stimulating and collaborative environment where students and researchers working in Artificial Intelligence and Data Science can interact on a daily basis. It offers a well-equipped and welcoming space designed to foster creativity, knowledge exchange, and teamwork. The lab has become a central hub for research activities, promoting close collaboration among PhD students, postdoctoral researchers, and faculty members.

The majority of the researchers associated with the Data Analytics Lab are actively involved in Evolutionary Computation and contribute to its advancement through high-impact publications, international collaborations, and participation in leading conferences. The environment is particularly supportive of young researchers, encouraging initiative, interdisciplinary thinking, and scientific ambition.

3. Description of the work to be carried out by the student

My research group is actively engaged in the development of advanced methodologies at the intersection of Genetic Programming (GP), machine learning, and optimization, with a particular focus on semantic-aware learning and interpretable symbolic modelling [1]. In this context, I propose a research plan that directly aligns with recent advances in transformer-based approaches for symbolic regression and semantic GP. In particular, the proposed work aims to build upon the emerging paradigm of integrating generative transformer models as semantic variation operators within evolutionary search, as recently introduced in [2]. This line of research represents a significant step forward in addressing one of the longstanding challenges in GP: guiding the search process in the semantic space of solutions while avoiding the uncontrolled growth in model complexity typically associated with geometric semantic operators. The main objective of the proposed research is to further investigate and extend transformer-based semantic operators for symbolic regression, with a focus on three key directions: (1) enhance the generalization capabilities of transformer-based variation operators by systematically studying the role of training data generation, including the diversity and representativeness of synthetic functions used during the model-building phase; (2) develop adaptive mechanisms to control the semantic distance between parent and offspring solutions during the evolutionary process, enabling a principled balance between exploration and exploitation in the semantic space; (3) explore the integration of these approaches within interpretable modeling frameworks. A further aspect of the research will be the extension of current approaches to more complex and heterogeneous data settings, including higher-dimensional problems and real-world applications where robustness and interpretability are critical. The goal is to design models that retain the key advantages of semantic GP, namely stability and effective search guidance, while exploiting the flexibility of transformer architectures. The expected outcome is a new class of hybrid methods that combine the strengths of Evolutionary Computation (EC) and deep generative modeling in a principled and scalable manner. The research environment at NOVA IMS provides an ideal setting for this work, offering strong expertise in EC, symbolic regression, and interpretable machine learning, as well as access to interdisciplinary collaborations and high-quality computational resources. Visiting students will benefit from close interaction with ongoing projects in these areas, contributing to high-impact publications and the advancement of state-of-the-art methodologies. I am confident that such a collaboration would be highly productive, fostering significant scientific contributions and strengthening international research cooperation in this rapidly evolving field.

[1] Vanneschi, L., Pereira, S. & Farinati, D. SLIM: the non-bloating genetic programming with geometric semantic mutations and meaningful semantic crossover. Genet Program Evolvable Mach 27, 9 (2026). https://doi.org/10.1007/s10710-026-09535-y

[2] Anthes, P, Sobania, D, and Rothlauf, F. Transformer Semantic Genetic Programming for Symbolic Regression. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '25). Association for Computing Machinery, New York, NY, USA, 952–960. 2025. https://doi.org/10.1145/3712256.3726412

4. Other relevant information

Lisbon, the capital and largest city of Portugal, stands as one of Europe's most vibrant and dynamic metropolitan centers. Recognized as an alpha-level global city, Lisbon plays a significant role in finance, commerce, media, entertainment, arts, international trade, education and tourism. Its unique ability to combine a rich historical heritage with a modern, forward-looking outlook makes it an especially attractive destination for students and researchers from around the world. The city offers a truly multifaceted environment, where tradition and innovation coexist harmoniously, creating fertile ground for intellectual growth, creativity, and cultural exchange. One of Lisbon's most distinctive features is its exceptional quality of life. The city is widely regarded as safe, welcoming, and inclusive, fostering a strong sense of community and well-being. This makes it an ideal setting for academic pursuits, as students can fully immerse themselves in their studies while enjoying a balanced and enriching lifestyle. Lisbon's cultural scene is both diverse and accessible, with museums, music, literature, gastronomy, and artistic events playing a central role in everyday life. This continuous exposure to culture not only enhances personal development but also contributes to a stimulating academic atmosphere. The city's mild climate further enhances its appeal. With long sunny periods throughout the year and moderate temperatures even in winter, Lisbon provides an excellent environment for both study and leisure. Its proximity to the Atlantic Ocean, as well as to natural parks and historic towns, offers numerous opportunities for outdoor activities and exploration. These characteristics have contributed to Lisbon's international recognition as one of the best cities to live in, including its ranking among the world's top cities by Monocle Magazine.

Within this outstanding urban context, the NOVA Information Management School (NOVA IMS) is strategically located in one of Lisbon's most prestigious and well-connected areas. The campus benefits from excellent public transportation links and immediate access to a wide range of services and facilities, ensuring convenience and accessibility for students and staff alike. Founded in 1989, NOVA IMS has established itself as a leading institution in Information Management and Data Science, currently hosting more than 3,000 students from over 80 countries. This highly international environment promotes diversity, intercultural dialogue, and the exchange of ideas, enriching both academic and personal experiences. NOVA IMS offers a comprehensive portfolio of programs, including undergraduate, postgraduate, master's, and doctoral degrees, all designed to meet the highest academic standards. NOVA IMS' mission is to play a pioneering and leading role in transforming data into value through Data Science. This is achieved through a strong and integrated commitment to teaching, research, and innovation, ensuring that knowledge creation is closely aligned with real-world impact. At the heart of NOVA IMS lies a vibrant and cohesive academic community composed of outstanding students, dedicated faculty, and highly skilled professional staff. The institution fosters an environment characterized by intellectual freedom, meritocracy, and collaboration, where individuals are encouraged to think critically, innovate, and pursue ambitious goals. This dynamic and supportive atmosphere empowers students and researchers to not only acquire knowledge but also to actively contribute to shaping the future of data-driven societies.